

Copyright

by

Jennifer Anne Durham-Fowler

2010

The Dissertation Committee for Jennifer Anne Durham-Fowler certifies that this is the approved version of the following dissertation:

Therapeutic Assessment with Couples

Committee:

Stephanie Rude, Supervisor

S. Natasha Beretvas

Stephen Finn

Alissa Sherry

Deborah Tharinger

Therapeutic Assessment with Couples

by

Jennifer Anne Durham-Fowler, B.A.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2010

Acknowledgements

I would like to thank my advisor, Stephanie Rude, for all of her guidance and encouragement throughout my years in graduate school. You modeled an intellectual curiosity and a love of research and ideas that I will carry with me throughout my career.

I would also like to thank Steve Finn, Pamela Schaber, and Marita Frackowiak at the Center for Therapeutic Assessment for donating so much of their time and energy to this project. This research would have been impossible without you, and I am forever grateful.

To the couples and therapists who participated in this study, I am so thankful to each of you for your commitment to a project that required such an investment of your time and emotional energy. I feel honored to have been given this window into a very personal part of your lives.

I would also like to thank my parents, Cathy and Ken, for instilling in me a love of reading, writing, and learning. To my mother, in particular, who once wrote a dissertation of her own while raising two teenage children, thank you for encouraging me and providing me with invaluable feedback as I applied to graduate school, wrote my prospectus and proposal, and, finally, this dissertation.

Finally, I would like to thank my husband Jeb, who gives me confidence and fills my life with laughter and surprises. You have been a wonderful companion throughout graduate school and I look forward to our next adventure together. I love you.

Therapeutic Assessment with Couples

Jennifer Anne Durham-Fowler, Ph.D.

The University of Texas at Austin, 2010

Supervisor: Stephanie Rude

Finn (2007) recently outlined procedures for applying Therapeutic Assessment (TA) techniques to work with couples. The current study used a time-series design to follow three heterosexual couples as they took part in a TA intervention. Participants were couples who were involved in ongoing couples therapy at the time of the study, but who felt they were not making satisfactory progress in therapy. Participants completed brief, daily measures of relationship satisfaction before, during, and after the TA. In addition, couples completed longer, standardized measures of relationship satisfaction, psychological symptomatology, and therapy progress. Qualitative feedback about the TA was also elicited from couples and their therapists. A time-series analysis revealed that all six participants reported significant improvement on at least some daily measures of relationship satisfaction, and that many of these improvements were sustained over a four-week follow-up period. In addition, four of the six participants reported fewer psychological symptoms at follow-up. Finally, qualitative feedback from participants revealed that all three couples and their couples therapists found the TA intervention to be a largely positive, useful experience.

Table of Contents

Chapter I: Introduction	1
Chapter II: Review of the literature	4
<i>Traditional (“information-gathering”) assessment models:</i>	4
<i>Therapeutic Assessment</i>	6
<i>Empirical research on collaborative and Therapeutic Assessment models</i>	10
<i>Why couples seek help</i>	12
<i>Couple interventions</i>	13
<i>Impasses in couples therapy</i>	15
<i>The use of psychological assessment in couples therapy</i>	16
<i>Does psychological assessment help couples?</i>	18
<i>Therapeutic Assessment with couples</i>	19
<i>The Consensus Rorschach</i>	20
<i>Intervention outcome research using single case designs</i>	22
<i>Time-series research</i>	23
<i>Goals of the current study</i>	24
Chapter III: Methodology	26
<i>Participants</i>	26
<i>Data collection procedure</i>	26
<i>Description of the TA intervention</i>	28
<i>Assessment instruments used in the TA intervention</i>	29
<i>Integrity checks</i>	33
<i>Measures</i>	33
<i>Couples’ daily measures</i>	34
<i>Standardized measures of relationship functioning and emotional wellbeing</i>	35
<i>Satisfaction with the Therapeutic Assessment intervention</i>	37
<i>Satisfaction with weekly couples therapy</i>	39
<i>Hypotheses and Analyses</i>	40
<i>Hypotheses</i>	40
Chapter IV: Results and Discussion	43
<i>Overview of results to be presented</i>	43
<i>Couple #1: Mary and Ray</i>	45
<i>Integrity rating of TA sessions</i>	46
<i>Results: Mary and Ray</i>	48
<i>Discussion: Mary and Ray</i>	78
<i>Couple #2: Susan and George</i>	80
<i>Integrity Rating of TA sessions</i>	81
<i>Results: Susan and George</i>	83
<i>Discussion: George and Susan</i>	110
<i>Couple #3: Elaine and Tom</i>	113
<i>Integrity rating of TA sessions</i>	114
<i>Table 4.5, cont</i>	115
<i>Results: Elaine and Tom</i>	116

<i>Discussion: Tom and Elaine</i>	139
Chapter V: Summary and Conclusions	141
<i>Summary of results from the three cases</i>	141
<i>Discussion</i>	145
<i>Strengths, Limitations, and Future Directions</i>	147
Appendix A: Demographic Questionnaire.....	150
Appendix B: Dyadic Adjustment Scale (DAS)	151
Appendix C: Marital Disaffection Scale (MDS)	154
Appendix D: Brief Symptom Inventory (BSI).....	155
Appendix E: Assessment Questionnaire-2 (AQ-2).....	157
Appendix F: Feedback Questionnaire—Client Form	159
Appendix G: Feedback Questionnaire—Therapist Form	160
Appendix I: Treatment Integrity Rating Forms	162
Appendix J: Means, standard deviations, and slopes for daily questions.....	165
References.....	175
VITA	183

Chapter I: Introduction

Psychologists have long used psychological tests as tools for gathering information about clients. Traditionally, information obtained from psychological assessments has been used “as a way to facilitate communication between professionals and to help make decisions about clients” (Finn & Tonsager, 1997, p. 378). In recent years, however, there has been a shift in the way many psychologists view the purpose of assessments, as well as the client’s role in the assessment process. Fischer (1970, 1972, 1985/1994) was one of the first to observe that when clients are invited to collaborate in their own assessments, they often appear to experience therapeutic benefits. Collaborative models of assessment have recently begun to grow in popularity due, in large part, to the work of Finn and his colleagues (Finn, 1996, 2007; Finn & Tonsager, 1992, 1997). Expanding on Fischer’s work, Finn developed Therapeutic Assessment (TA), a semi-structured technique for conducting collaborative assessments that incorporates elements of both psychotherapy and psychological testing.

Although research on the therapeutic value of TA is still in its early stages, initial results look promising. In studies of college students seeking services at counseling centers, TA has been associated with decreased psychological symptomatology, increased self-esteem, and an increased sense of hopefulness (Finn & Tonsager, 1992; Newman & Greenway, 1997). Other researchers have found that clients who participate in TA before beginning a course of psychotherapy are likely to be more committed to therapy and experience a stronger alliance to the therapist than clients who participate in more traditional assessments (Ackerman, Hilsenroth, Baity, & Blagys, 2000; Hilsenroth, Peters, & Ackerman, 2004). In addition, preliminary studies of TA as a therapeutic intervention for children and families have found TA to be associated with decreased psychological symptomatology in children and improved family functioning (Smith, Wolf, Handler, and Nash, 2009; Tharinger, Finn, Gentry, Fowler, et al., 2009).

Finn (2007) recently suggested TA may also be a useful intervention for couples who are experiencing distress or dissatisfaction in their relationships. In a recent survey, Doss and colleagues (2004) found that couples seek marital therapy with a variety of

presenting concerns, including problematic communication, lack of affection, and the goal of improving the relationship for the sake of the children. In addition, these authors found spouses rarely agree with one another about reasons for seeking therapy and may not always clearly express their presenting concerns to the therapist. Although outcome research has shown that a number of different treatment modalities may be useful in reducing relationship distress, Doss et al. suggested therapeutic outcomes may suffer when therapists and clients lack insight into the complex issues and agendas that each member of the couple brings to the table. Given these findings, it stands to reason that both couples and the therapists who work with them might benefit a great deal from psychological assessments that collaboratively explore ways in which two personalities come together to function (or not function) as a unit.

Finn (2007) recently outlined procedures for applying TA techniques to work with couples. During TA, the assessor administers psychological tests, such as the MMPI-2 and Rorschach, and then works collaboratively with the couple (and often the couple's therapist) to answer the couple's questions about themselves and their relationship. Couples and their therapists can also use the TA to help resolve impasses that may have occurred in couples therapy. Although initial case material (Finn, 2007) has suggested TA may be useful in helping couples improve relationship satisfaction and make greater strides in therapy, empirical research is needed to examine the clinical utility of TA with couples.

The current study used a time series design to follow three heterosexual couples as they took part in a TA intervention. The participants were married or cohabitating couples who were involved in ongoing couples therapy at the time of the study, but who felt they were "stuck" or not making satisfactory progress in therapy. The assessors conducting the TA interventions were three psychologists who were in practice at the Center for Therapeutic Assessment in Austin, Texas. Participants completed brief, daily measures of relationship satisfaction throughout three phases of the study: a pre-intervention baseline phase (duration: four weeks), an intervention phase (duration: six to seven weeks), and a post-intervention follow-up phase, which began one month after the

couples' final TA session (duration: four weeks). In addition, the couples and their therapists completed longer, standardized measures of relationship satisfaction, symptomatology (psychological distress), and therapy progress at three time points: pre-intervention, post-intervention, and one-month follow-up. It was hypothesized that the TA would lead to an increase in relationship satisfaction and a decrease in variables such as disaffection and interpersonal conflict. It was also hypothesized that participants would report decreases in psychological symptoms. Participants were also expected to report feeling satisfied with the TA intervention. Finally, it was hypothesized that, after the TA intervention, both the couples and their therapists would report an increase in their satisfaction with the progress being made in couples therapy. A time series analysis was used to test whether there was a significant change in daily measures of relationship satisfaction between pre- and post-intervention (Borckardt, Nash, Murphy, Moor, Shaw, & O'Neil, 2008). Visual analysis was also used to examine whether couples reported changes on global measures of relationship satisfaction, psychological symptoms, and therapy progress (Kazdin, 1983).

Chapter II: Review of the literature

Traditional (“information-gathering”) assessment models:

Psychological assessments can serve a number of purposes. Therapists and psychiatrists may request assessments as part of diagnostic clarification and treatment planning. Schools may make assessment referrals in order to determine whether students meet criteria for disability or special education services. Lawyers and judges use psychological test results to inform their decisions about issues such as child custody or competency to stand trial. In addition, many people seek assessments with the simple goal of learning more about themselves.

Finn and Tonsager (1997) described the traditional approach to psychological assessment as an “information-gathering” model. In this model, assessors administer standardized tests, interpret test results, and make recommendations based on these interpretations. Although assessors using this model often take steps to establish good rapport with clients, they tend to keep social interactions with clients to a minimum and avoid allowing their own subjective experience of clients to influence their interpretation of test results. The results of psychological tests are often used to describe clients “in terms of already existing categories and dimensions (e.g., *schizophrenic, IQ of 100, 2-7 code type on the MMPI-2*)” (p. 378). Test findings are presented to clients through verbal or written feedback, and clients are usually not encouraged to question these results.

Although the information-gathering model provides an often valuable and efficient means of summarizing a large amount of information about clients, it also has several limitations. The assessment process can be quite mysterious and anxiety-provoking, especially for clients who are given little information about or rationale for the tests they are taking. While assessment reports are often sent to the clients’ therapists, psychiatrists, teachers, employers, and lawyers, it is not uncommon for clients to undergo an assessment without ever seeing a copy of their report. In their survey of 719 psychologists who regularly conduct neuropsychological and personality assessments, Smith, Wiggins, and Gorske (2007) found that although 71% of respondents reported frequently giving in-person feedback to clients, only 47% reported frequently providing

their clients with written reports. Indeed, historically, many psychologists were trained *never* to show an assessment report to a client, as the information in the report might be confusing or upsetting to the client, and could even provoke some dissatisfied clients to litigation (Fischer, 1972). The American Psychological Association's (APA; 2002) code of ethics now states psychologists must take "reasonable steps" to explain test results to clients. However, when assessment feedback takes the form of an "expert" delivering "facts" to a patient in either a report written in clinical language or a feedback session that does not involve a dialogue between assessor and client, it is likely that many clients may come away from their assessments feeling vulnerable, disappointed, or confused.

Collaborative/individualized assessment

In recent years there has been a shift in the way many psychologists view the client's role in the assessment process. Increasingly, assessors are beginning to invite clients to collaborate in their own assessments. While an assessor using the information-gathering model of assessment might be thought of as "a scientist examining a sample through the lens of a microscope," an assessor using a collaborative model takes on the role of a "participant-observer who play[s] an active, influential role in shaping the assessment process, along with [her] clients" (Finn & Tonsager, 1997, p. 379). Constance Fischer (1970, 1972, 1985/1994) was one of the first to advocate this new approach to psychological testing. Fischer's approach, rooted in the human science tradition, called for assessors to view clients as "co-evaluators," as "it is the client himself who is in the best position to confirm or clarify the evaluator's impressions" (Fischer, 1970, p. 71). Fischer urged assessors to "dialogue rather than diagnose," encouraging clients to share their own impressions of test results. She also recommended writing reports in "everyday language," free from clinical jargon, and allowing clients to critique and amend the assessor's interpretations of test results. In addition, Fischer asserted that clients should always be allowed to have the final say over who would ultimately read the report, and under what conditions.

Finn and Tonsager (1997) noted that information-gathering and collaborative/therapeutic models of assessment need not be mutually exclusive. Collaborative assessment can be used for many of the same purposes as information-gathering assessment (e.g., treatment planning, educational placement), just as information-gathering assessment can have therapeutic benefit. In fact, clinicians and researchers have long noted that the simple act of providing clients with feedback may have therapeutic value (e.g., Butcher, 1990; Dorr, 1981; Fischer, 1970). For example, Allen, Montgomery, Tubman, Frazier, and Escobar (2003) found that college students who took part in a 15-minute, personalized feedback session after taking the Millon Index of Personality Styles (MIPS; Millon, Weiss, Millon, & Davis, 1994) reported higher levels of self-understanding, self-liking, self-competence, and global self-esteem than controls who took the test but did not receive feedback. In addition, participants in the personalized feedback condition reported having better rapport with the examiner than controls.

Therapeutic Assessment

The popularity of collaborative models of assessment has grown due, in large part, to the work of Finn and his colleagues (Finn, 1996, 2007; Finn & Tonsager, 1992, 1997). As collaborative and therapeutic assessment models are beginning to take on many different forms, it is important to distinguish between various terms used to describe interventions that fall under the umbrella of “therapeutic assessment.” Finn (2007) uses the term “therapeutic assessment” (lowercase) to describe any assessment that is “based on the intent to use psychological assessment to help patients directly, rather than just indirectly, as with traditional assessment” (p. 5). Therapeutic assessment can be thought of as an attitude or approach to assessment, rather than a set of specific guidelines for conducting assessments. Finn uses the term “collaborative assessment” to refer to a more specific type of therapeutic assessment. Collaborative assessment involves a comprehensive effort to engage the client in multiple phases of the assessment process—including (a) framing the reasons for the assessment, (b)

observing test responses and behaviors, (c) discovering the significance of those responses and behaviors, (d) coming up with useful recommendations, and (e) drafting summary documents at the end (p. 5).

Finally, the term “Therapeutic Assessment” (TA; uppercase) is used to describe Finn’s (2007) semi-structured technique for conducting collaborative assessments. This more structured technique will be the focus of this dissertation.

An important component of TA is the presenting of assessment results in order of ego-syntonicity—that is, the degree to which feedback is consistent or discrepant with what the client already thinks about himself. Finn (2007) suggested assessors begin summary or feedback sessions by presenting clients with findings that are congruent with the way in which they already perceive themselves (Level 1 findings). Next, assessors should discuss findings that “reframe or amplify clients’ usual ways of thinking about themselves” (Level 2 findings, p. 8). Level 2 findings will likely lead a client to think about himself in a slightly different way without threatening his sense of identity or self-worth. Finally, if the summary session seems to be going smoothly, the assessor can introduce Level 3 findings. These findings will likely be at odds with a client’s previously held self-views. Although clients may initially reject Level 3 findings, Finn has noted that in many cases, clients have reported assimilating at least some of the Level 3 findings in the weeks and months following the assessment. This method of presenting feedback is consistent with Swann’s (1997) self-verification theory, which posits that people are more willing to take in information that confirms their preexisting self-views or schemas (even when this information is negative), and less willing to accept feedback that contradicts these views. Indeed, research by Schroeder, Hahn, Finn, and Swann (1993) indicated clients are most likely to view assessment feedback as being accurate and useful when it is presented beginning with Level 1 feedback and ending with Level 2 feedback.

Finn (2007) outlined six basic steps that comprise a TA. It should be noted that although these are the steps that make up a *typical* TA intervention, they can be modified depending on the assessment goals, as well as clients’ time or financial constraints. In

Step 1, the assessor meets with the client(s) for an initial session in which clients are asked “what puzzles, questions, or quandaries they [have] about themselves” (Finn, 2007, p. 10). The assessor asks clients to share their “best guesses” about the answers to these questions. Clients are also invited to revise and amend their list of assessment questions throughout the assessment process. Some of the goals of this first session include establishing rapport, helping clients become curious about their problems, and orienting them to the collaborative nature of the assessment.

In Step 2, clients take part in one or more standardized testing sessions. Tests are chosen based on the clients’ presenting concerns and may include cognitive tests (e.g., WAIS-III, Bender-Gestalt Test), projective personality tests (e.g., Rorschach, TAT), and self-report measures of emotional and psychological functioning (e.g., MMPI-2). In many ways, Step 2 resembles traditional psychological assessment. However, in TA, the assessor pays close attention to the way in which the client is experiencing the assessment process. For example, the assessor will provide the client with rationale for using tests that are less face-valid, such as the Rorschach and MMPI-2. The assessor will also ask the client to talk about his or her reactions to taking each test. Asking clients to talk about their experiences in this way not only provides important information about the clients, but also helps to decrease their anxiety and strengthen the working alliance with the assessor.

In Step 3, clients take part in what Finn (2007) calls an “assessment intervention session.” The goal of this session is “to bring into the room those problems-in-living of the client that are the focus of the assessment, where they may be observed, explored, and addressed with various therapeutic interventions” (p. 14). At this point in the assessment process, the assessor will have scored the standardized tests and will have formulated hypotheses about possible answers to the client’s assessment questions. Given that some feedback (particularly Level 3 feedback) will be difficult for the client to integrate into her preexisting understanding of herself, the assessor can use the assessment intervention session as an opportunity to make this feedback more palatable to the client by eliciting her problem behavior in vivo and then exploring possible real-life solutions to the

problem. Finn suggested three strategies for eliciting problem behavior: using “(a) standardized tests you haven’t yet used (perhaps in unstandardized ways); (b) role plays, art projects, psychodrama, or other less-structured activities; or (c) nonstandardized, or even out-of-date assessment materials that you will not score or attempt to interpret by nomothetic norms” (p. 89).

To illustrate the third strategy for conducting assessment intervention sessions, Finn (2007) described the case of a client whose assessment question was, “Why can’t I succeed at anything?” In this case, the assessor, after scoring the client’s IQ test (WAIS-III), academic achievement test (WIAT-II), and personality inventory (MMPI-2), suspected that although the client was very intelligent, his low self-esteem was causing him to underestimate his abilities and often sabotage his own chances at success in school, work, and other areas of his life. In an attempt to bring the client’s problem to life during the assessment intervention session, the assessor administered the Bender Visual Motor Gestalt Test (BVMGT; Bender, 1938). As was predicted, the client attempted to give up early, but with encouragement from the assessor, he completed the test and his performance was nearly perfect. The client, however, believing he had failed horribly, was visibly upset, calling himself “stupid.” When the examiner explained that the client had, in fact, done quite well on the test, the client was quite surprised and confused. This led to a conversation in which the client and examiner discussed ways in which the client’s low self-confidence, rather than lack of ability, seemed to keep him from succeeding in life. They then discussed contexts in which this problem was likely to be most pronounced and explored possible solutions to the problem. Finn noted that when clients’ problems are brought into the assessment room in this way, they are much more likely to internalize and make use of feedback than they would if the information were simply presented to them in a written report. In addition, this session allows clients to begin brainstorming about real-life solutions shortly after encountering a problem.

In Step 4 of TA, clients take part in a summary and discussion session with the assessor. In this session, the assessor begins by discussing Level 1 findings, then works up to Level 2 and possibly Level 3 findings. During this step, the assessor strives to be

emotionally supportive as clients begin to take in the new information. The assessor also encourages clients to take an active role in the discussion, inviting them to “agree, disagree, revise, and give real-life examples” (Finn, 2007, p. 10). In Step 5 the assessor provides written feedback to clients in the form of a letter. This letter incorporates clients’ own interpretations and contributions from the summary and discussion session. In addition, clients are invited to make their own revisions to drafts of the letter. Finally, in Step 6, clients attend a follow-up session approximately 2-3 months after their last meeting with the assessor. The purpose of this session is to give clients an opportunity to discuss new questions or concerns that have arisen since the last session.

Empirical research on collaborative and Therapeutic Assessment models

Although research on the clinical utility of TA is still relatively new, results from initial studies look promising. In a seminal study, Finn and Tonsager (1992) investigated the therapeutic value of using a collaborative assessment model to give MMPI-2 feedback to college students. The participants were 60 students awaiting outpatient therapy at a university counseling center. The 32 participants in the experimental condition participated in two sessions: an assessment session and a feedback session. The assessment session began with a 30-minute clinical interview. Through collaborative discussion, the examiner solicited assessment questions from each participant (e.g., What would the participant like to get out of the assessment?). At the end of the session, participants completed the MMPI-2, as well as several dependent measures, which included self-report measures of self-esteem and general psychological distress. Two weeks later (Time 2), the examiner discussed each participant’s MMPI-2 results with the participant in a collaborative feedback session. Participants were encouraged to take an active role in this session, sharing their reactions to the examiner’s feedback and altering or amending the examiner’s hypotheses. Participants completed dependent measures immediately following this feedback session, and again two weeks later (Time 3). The 28 participants in the control condition participated in a 30-minute interview in which they

talked about their concerns and completed research measures (but not the MMPI-2). Two weeks later (Time 2), they participated in another 30-minute interview in which they continued to discuss their concerns. They completed the dependent measures immediately following this session, and again two weeks later (Time 3). At Time 2, participants in both conditions also completed a measure in which they reported on their subjective impressions of the sessions.

The authors found participants in the experimental condition showed a significant decrease in symptomatic distress over time, and fewer symptoms at Time 3 when compared to controls. Participants in the experimental group also showed higher self-esteem than controls at both Time 2 and 3, and reported feeling more hopeful than controls. Finally, participants who reported having more positive feelings about the assessment experience were more likely to experience increases in self-esteem and decreases in symptomatology. Newman and Greenway (1997) replicated the first two findings from Finn and Tonsager's study in a sample of 60 Australian university students.

Results from other studies have suggested collaborative assessment, when conducted prior to psychotherapy, may help increase clients' commitment to or willingness to participate in therapy. Ackerman, Hilsenroth, Baity, and Blagys (2000) found that when clients at a community clinic took part in a collaborative assessment prior to beginning formal psychotherapy, they were much less likely to ultimately terminate therapy against medical advice than clients who participated in a more traditional, information gathering model of assessment. In addition, the results presented in Ackerman et al.'s (2000) article, as well as later results from the same study (Hilsenroth, Peters, & Ackerman, 2004), indicated collaborative assessment may be associated with higher levels of client-therapist alliance during subsequent psychotherapy when compared with alliance levels reported by clients who received traditional assessment. Similarly, Hanson, Claiborn, and Kerr (1997) found that college students participating in a brief "interactive" (collaborative) assessment were more likely to rate their sessions "deep" and their counselors as expert, trustworthy, and attractive, than

students participating in an assessment that involved unilateral feedback from the assessor.

Recent studies are also beginning to lend empirical support to TA as an effective intervention for children and families. Tharinger, Finn, Gentry, Hamilton, Fowler, Matson, Krumholz, and Walkowiak (2009) found TA with children to be associated with significantly decreased child symptomatology and improved family functioning. Similarly, in a recent single-case study using time-series analysis, Smith, Wolf, Handler, and Nash (2009) found that TA with a family was associated with a decrease in problematic behaviors in a 9-year-old boy.

To date, all of the empirical research on collaborative and Therapeutic Assessment has focused on individual (Ackerman et al., 2000; Finn & Tonsager, 1992; Hilsenroth et al., 2004; Newman & Greenway, 1997) and family clients (Smith et al., 2009; Tharinger et al., 2009; Tharinger, Finn, Wilkinson, & Schaber, 2007). However, an accumulating body of case material suggests collaborative and Therapeutic Assessment techniques may also be useful for psychologists working with couples, either as a precursor or supplement to traditional couples therapy (Butcher, 1990; Dorr, 1981; Finn, 2007; Richman & Davidoff, 1971). In the following section, I will review the literature on common problems that lead couples to seek therapy and assessment. I will also review the literature on some of the most commonly-used couples therapy interventions and their efficacy. Finally, I will provide an overview of the ways in which psychologists have traditionally used psychological assessment with couples.

Why couples seek help

In their survey of 147 married couples seeking couples therapy, Doss, Simpson, and Christensen (2004) found couples seek marital therapy with a variety of presenting concerns, with the two most commonly reported concerns being problematic communication and lack of affection. Other oft-cited reasons for seeking treatment include concerns related to divorce or separation, desire to improve the relationship, concerns about arguments or anger, and wanting to improve the relationship for the sake

of the children. In addition, Doss and colleagues found spouses rarely agree with one another about reasons for seeking therapy and may not always clearly express their presenting concerns to therapists. Given these findings, one might conclude that therapeutic outcomes are likely to suffer when both therapists and clients lack insight into the complex issues and agendas that each member of the couple brings to the table.

The consequences of marital discord can be far-reaching. Research has consistently shown that conflict between parents is likely to have a negative impact on children's emotional, cognitive, academic, and social functioning (see Cummings & Davies, 2002 for a review). In addition, marital discord may also play a causal role in the development of Major Depressive Disorder (Whisman & Bruce, 1999) and alcohol use disorders (Whisman, Uebelacker, & Bruce, 2006) among both women and men. Marital dissatisfaction has also been associated with DSM-IV diagnoses such as Bipolar Disorder, Posttraumatic Stress Disorder, and Generalized Anxiety Disorder, although the nature of the relationship is unclear in the absence of longitudinal data (Whisman, 2007). Finally, the direct and indirect consequences of divorce are estimated to cost taxpayers several billion dollars every year (Schramm, 2006).

Couple interventions

Given that marital problems are likely to affect the lives of both individuals and families in a variety of ways, research identifying effective interventions is crucial. In this section, I will summarize the efficacy research on couples interventions and describe two of the most commonly-used interventions. Over the years, numerous empirical studies have examined the efficacy of various couples interventions. In their meta-analyses of outcome studies of marital and family therapy, Shadish and colleagues (Shadish & Baldwin, 2003; Shadish, Ragsdale, Glaser, & Montgomery, 1995) have found that, in general, couples therapy is more effective than no therapy. Shadish et al. (1995) found that in studies where the presenting problem was global marital dissatisfaction ($n=16$), the average effect size was .71. In studies where the presenting problem was related to communication or specific problem-solving complaints ($n=7$), the average effect size was

.52. In addition, Shadish et al. found that no one treatment modality clearly outperformed any other. In their meta-analysis of randomized controlled trials of behavioral marital therapy (BMT; n=30), Shadish and Baldwin (2005) found that BMT was consistently more effective than no treatment, with an average effect size of .585. In another recent meta-analysis of couples therapy outcome studies (n=23), Wood, Crane, Schaalje, and Law (2005) found that although both BMT and emotion focused therapy (EFT) were associated with significant improvement in marital adjustment across studies, EFT appeared to be significantly more effective than BMT in studies of couples reporting moderate levels of marital distress.

Two of the most commonly studied interventions for couples are behavioral marital therapy (Jacobson & Margolin, 1979) and emotion focused therapy (Greenberg & Johnson, 1988; Johnson, 1996). Rooted in social learning theory, the primary goal of behavioral marital therapy (BMT) is to teach couples the skills needed to sustain a mutually satisfying, long-term relationship (Jacobson, 1981). BMT therapists focus on helping clients develop effective problem-solving and communication skills. When teaching communication skills to couples, therapists focus on the resolution of conflict, placing less emphasis on emotional expression. Therapists teach new communication skills through instruction and modeling, and through behavior rehearsals in which couples practice new communication skills and receive feedback from the therapist. In addition, couples are taught to use behavior reinforcement strategies to increase the frequency of desired behaviors from one another. For example, a wife might learn ways of reinforcing specific behaviors in her husband, such as doing household chores or asking her about her day. Throughout BMT, therapists emphasize couples' strengths and attempt to increase positive behaviors, rather than focus on problematic behaviors.

Emotion focused therapy (EFT), developed by Greenberg and Johnson (Greenberg & Johnson, 1988; Johnson, 1996), has its roots in systems theory, attachment theory, and humanistic/experiential perspectives. It is a brief model, with the course of therapy lasting approximately 8 to 15 sessions. It is also a constructivistic approach in which clients' problems are conceptualized in terms of the social context from which

they arose, rather than as symptoms of inherent pathology. Greenberg and Johnson (1988) outlined four general principles of EFT with couples: First, therapists focus on clients' present experience and tend to shift the focus away from discussions of each partner's past experience. For example, if a couple comes to a session wanting to discuss a recent argument, the therapist will encourage the couple to not only discuss the argument, but to re-experience any aspects of the argument that remain unresolved. Second, EFT therapists focus on "primary emotional experiences, especially each partner's experience of vulnerability and/or fear, and [have] partners communicate these underlying experiences to each other" (p. 58). Third, EFT therapy focuses on "the interactive process rather than problematic issues...that is, how the couple fights is more important than what they fight about" (p. 58). The therapist avoids trying to solve particular problems for a couple, and instead helps them recognize problematic patterns of interaction that occur in their relationship. Fourth, therapists focus on "restructuring the interaction using the newly accessed primary emotions to motivate new behavior" (p. 58). The underlying assumption is that once clients have accessed their primary emotions, they will be more willing and able to express their true needs and wants to their partners.

Impasses in couples therapy

As with any type of psychotherapy, there are times when couples and their therapists find themselves feeling stuck or frustrated with the progress being made in treatment. In Stewart and Chambless's (2008) survey of 591 private practice therapists, respondents reported working with clients for a median of 12 sessions before concluding that therapy was not making progress. Eighty-five percent of therapists in the sample reported consulting with colleagues when working with clients who were not progressing. Although this study did not provide information about the types of consultation sought by therapists, it is clearly not uncommon for therapists to turn to their colleagues for help when they feel they are at an impasse with clients.

Although therapeutic impasses are not uncommon in clinical work, little has been written about the possible causes of impasses in couples therapy. In one exception,

Johnson, Makinen, and Millikin (2001) theorized that “attachment injuries” in one or both members of a couple may lead to impasses in couples therapy. These authors defined an attachment injury as “a wound that occurs when one partner fails to respond to the other in a critical time of need, and this incident becomes a clinically recurring theme and creates a barrier to relationship repair” (p. 153-154). To the outside observer, attachment injuries can be seemingly large or small, and may include obvious betrayals, such as infidelity, or more subtle disappointments, such as the failure of one partner to respond to the other partner’s emotional needs during a time of distress. Johnson and colleagues suggested that if these injuries are not brought to light and properly processed in therapy, the injured partner or partners may not be able to let go of feelings of disappointment and mistrust.

Whether impasses occur due to attachment injuries or other factors, it is clear that every couples therapist is likely to feel stuck from time to time in his or her clinical work. At these times, psychological assessments may be useful in helping therapists understand more fully their clients’ problematic dynamics and develop new strategies for becoming “unstuck” in therapy. In the following sections, I will review the ways in which assessment has been traditionally used as a precursor or supplement to couples therapy and will then discuss how Therapeutic Assessment can be used to help couples and their therapists resolve impasses.

The use of psychological assessment in couples therapy

To date, most of the writing about couples assessments has focused on using psychological assessment findings to inform treatment decisions. Floyd, Haynes, and Kelly (1997) identified five pre-treatment assessment goals for therapists working with couples:

- (a) specifying both problems and strengths for the couple; then (b) identifying variables that maintain and mediate problems and prevent couples from employing a new, more effective behavioral repertoire...(c) specifying environmental resources that might affect treatment outcomes; (d) establishing

treatment goals; and (e) identifying reinforcers that can be used in treatment to achieve these goals (p. 350).

Toward this end, Floyd and colleagues recommended using various self-report and observational measures of relationship functioning, such as the Dyadic Adjustment Scale (Spanier, 1976) and the Marital Interaction Coding System (Weiss & Tolman, 1990).

Other authors recommend collecting data about each partner's individual level of psychological and emotional functioning prior to couples therapy. For example, Nurse (1999) suggested conducting a brief assessment using standardized personality measures, such as the MMPI-2 or MCMI-III, at the onset of couples therapy. Information derived from these tests can give therapists valuable information about ways in which each partner's personality style or psychological functioning may be contributing to the couple's current problems. In addition, if these measures reveal serious psychological problems, such as suicidal ideation or a personality disorder, in one or both partners, the therapist may choose to recommend individual therapy or medication evaluation in addition to, or in lieu of couples therapy.

Some researchers and therapists also advocate using projective measures, such as the Rorschach Inkblot Test (Exner, 1993), in order to gain a deeper understanding of the ways in which individual-level variables affect a couple's relationship (Dorr, 1981; Nurse, 1999). Although self-report measures can provide valuable information about each partner's psychology, projective tests have the added benefit of identifying potential problems (e.g., depression, poor reality testing) that may not be in clients' conscious awareness, or that clients may have downplayed on self-report measures.

Although psychological assessment is most often described as a tool to be used prior to treatment, there are several contexts in which it may be useful for therapists and clients who have already begun treatment. Richman and Davidoff (1971) listed four conditions under which psychological testing can be an especially useful component of couples therapy:

First, when there has been a shift in the marital equilibrium so that the old defenses do not work and a crisis results; second, when an impasse has been

reached in therapy, resulting in or stemming from heightened anxiety and tension; third, when the therapist suspects the presence of some covert pathological maneuvers and resistances which are undermining the treatment; and fourth, when blaming and scapegoating has taken over, and the marital interaction has become increasingly destructive (p. 440).

Does psychological assessment help couples?

To date, there has been little research on the therapeutic value of psychological assessment with couples. However, the small amount of case material that does exist suggests assessment may contribute to positive psychotherapy outcomes. Richman and Davidoff (1971) reported five clinical cases in which psychological assessment was used at either the beginning or middle stages of marital therapy. In each case, clients were administered individual, standardized tests, as well as “interaction” tests, in which both members of the couple were asked to agree on responses to tests, such as the Rorschach. The authors reported that after clients were given face-to-face feedback from an assessor who interpreted test scores in “non technical language,” all five couples experienced improvements in relationship satisfaction and /or an increased willingness to engage in marital therapy.

Findings from another, more recent study suggest assessment alone may help improve relationship satisfaction. Worthington and colleagues (1995) examined whether receiving feedback after a couples assessment would improve relationship satisfaction. These authors randomly assigned 48 college student couples to either an “assessment-feedback” condition or an “assessment only” condition. The 28 couples in the assessment-feedback condition attended two assessment sessions in which they completed several measures of relationship adjustment, discussed their relationship with the assessor, and participated in a videotaped behavioral observation task. The following week, each couple met with the assessor and received both oral and written feedback about their relationship. The 20 couples in the assessment only condition completed the same measures of relationship adjustment, but did not meet with an assessor or receive

feedback. The authors found that couples in the assessment-feedback condition made small, significant gains in their relationship satisfaction between pre-and post-assessment, while the couples in the assessment only condition did not. In addition, participants in the assessment-feedback reported a significant increase in dedication to the relationship between pre- and post-assessment, while couples in the assessment only condition reported no gains in dedication.

Therapeutic Assessment with couples

Finn (2007) recently outlined steps for conducting TA with couples. This format is similar in many ways to TA with individuals. If the couple is already in therapy, the assessor will usually speak with the couple's therapist prior to the assessment, either in person or over the phone, to discuss the couple's presenting concerns, as well as the therapist's questions for the assessment. The assessor will then meet with the couple for an initial interview in which the couple discusses their concerns and develops assessment questions. In the following sessions, the assessor meets with each member of the couple separately for a one-on-one interview in order to obtain more information about each partner. Assessment questions in couples TA might include both individual-level questions (e.g., What can I do to help myself feel less anxious?) and relationship-level questions (e.g., Why isn't therapy helping us feel closer to each other?). After these initial interviews, each partner undergoes psychological testing. The tests used will depend on the couple's presenting problems, but often include the Rorschach and MMPI-2. After each partner has completed individual testing, the couple meets with the assessor for an assessment intervention session, which often includes a Consensus Rorschach (to be described in the following section). Finally, the couple meets with the assessor for a summary/discussion session in which the findings from the assessment are discussed. Often, the couple's therapist will attend this session so that he or she can both emotionally support the couple and hear feedback as it was presented to the couple. In addition, the assessor may also choose to meet privately with the couple's therapist to discuss ways in which assessment findings might affect treatment planning. Soon after

the final session, the assessor sends the couple a letter summarizing the assessment findings. In many cases, the couple is invited to attend a follow-up session several months later in order to discuss the couple's progress and address any new or unanswered questions from the assessment.

Finn (2007) presented clinical material from several cases in which couples benefited from TA. In one case, for example, a husband and wife who had been in marital therapy for years reported feeling distant and frustrated with one another. Both partners attributed many of their relationship problems to the wife's longstanding depression. Among other things, the TA revealed that the husband had also been struggling with a deep and unacknowledged depression, and had never known how to ask for help and support from his wife. This new information helped bring balance to the relationship, with the wife becoming more empathic toward her husband and the husband becoming more comfortable in asking his wife for emotional support. One month after the assessment, the wife also reported feeling much less depressed and both partners reported feeling better about their relationship.

The Consensus Rorschach

The "Consensus Rorschach," is a projective and interactive task that is often used in TA as part of a couple's assessment intervention session (Finn, 2007). In this procedure, the examiner presents Rorschach cards to the couple and asks them to come up with responses that they can both see and agree upon. Asking a couple to provide joint responses in this way gives the assessor a unique opportunity to observe firsthand the way in which a couple interacts with one another when asked to work together on an ambiguous and often emotionally-charged task. Numerous Consensus Rorschach techniques have been described over the years (e.g., Aranow, Reznikoff, & Moreland, 1994; Handler, 1997; Noy-Sharav, 2005). Many assessors will administer a full Rorschach to the couple and then score the responses using Exner's (1993) Comprehensive System. As this procedure can take up to four hours, however, some authors (e.g., Finn, 2007; Noy-Sharav, 2005) have suggested shortening the consensus

Rorschach by only presenting the couple with a select few cards. I will now describe this shorter version of the consensus Rorschach in greater detail.

When using the Consensus Rorschach in an assessment intervention session with a couple, Finn (2007) suggested beginning the task by giving the couple the following instructions:

I have a series of cards here that have inkblots on them. As I show you each card, I want you to talk out loud together about what you see. Your task is to come up with responses that both of you can see and both of you agree upon reporting. When you agree, let me know and I will write down your response and ask you some questions about it (p. 160).

The assessor then presents the couple with cards I, II, IV, VII, and X. Throughout the task, the assessor observes the couple's interactions, making notes about "significant events and patterns" (Finn, 2007, p. 160). After the couple has finished, there is a short break during which the assessor looks over his notes and begins to think about ways in which his observations from the Consensus Rorschach relate to the couple's assessment questions. After the break, the assessor brings the couple back into the room and invites them to share their thoughts and feelings about the task they just completed. In some cases, the examiner then shows the couple clips from a videotape of the Consensus Rorschach administration, asking open-ended questions about what the couple notices as they watch themselves interacting. The assessor then shares his observations, highlighting "especially systemic aspects of the interactions, for example, 'Did you see how you both play a part in the pattern we saw?' 'John, what did you notice about your response when Mary did X?' 'Mary, did you notice what you did right after that?'" (p. 162). Next, the assessor asks the couple to think of a strategy they might have used during the task to minimize problematic interactions. Using one of the Rorschach cards that was not administered the first time around, the assessor then asks the couple to try out this new approach to the task. If, after practicing in the room with the assessor, the new approach is successful, the assessor can ask the couple to think about ways in which they might use this new strategy in their daily lives.

Intervention outcome research using single case designs

In the eyes of many researchers in the fields of counseling and clinical psychology, randomized controlled trials (RCTs) are the gold standard for determining whether psychotherapy interventions actually work. However, due to the substantial commitment of time and money required by large n RCTs, they are not always a practical or feasible option for researchers interested in conducting outcome research. In addition, while RCTs provide valuable information about the overall effectiveness and generalizability of different interventions, they often obscure other important aspects of what happens during a therapy intervention and why it works.

An advantage of using single case designs is that, unlike large n studies, they allow the researcher to closely and rigorously examine the ins and outs of a single case. Kazdin (1983) noted that a defining feature of clinical psychology is its focus on the individual. Indeed, throughout the history of clinical and counseling psychology, the study of individual cases has informed the development of psychotherapy techniques. Findings from small n and single case studies are often dismissed as being merely “anecdotal” evidence of treatment efficacy; however, researchers are increasingly calling attention to the potential of single case designs for contributing to the empirical literature on treatment effectiveness (Borckardt et al, 2008; Chambless & Hollon, 1998; Chambless & Ollendick, 2001; Kazdin, 1983). The American Psychological Association’s (APA’s) Task Force on Evidence-Based Practice (2006) noted that single-case experiments “are particularly useful for establishing causal relationships in the context of an individual” (p. 274). In addition, APA’s Division 12 Task Force on Promotion and Dissemination of Psychological Procedures has stated that a psychological intervention can be considered “probably efficacious” once a small series of well-designed, single-case experiments have demonstrated its therapeutic benefits (Chambless & Ollendick, 2001).

Time-series research

One type of single-case design, which allows for the systematic study of an intervention's efficacy, is the time-series study. While many single-case designs rely solely on visual analysis of data to look for changes in the dependent variable (e.g., symptoms or depression or anxiety), time-series designs allow researchers to statistically test whether an intervention is associated with significant changes in the level and/or trend of the dependent variable. While visual analysis is always an important component of single-case research, Kazdin (1983) notes time-series designs can provide useful additional information when visual analysis yields ambiguous results; for example, "when there is a trend in the therapeutic direction in baseline, when variability is large, or when treatment effects are neither rapid nor marked" (p. 250).

Borckardt and colleagues (2008) recently outlined procedures for using time-series analysis in psychotherapy research. The time-series studies described by these authors typically contain three phases: (1) a pretreatment, baseline phase, (2) an intervention phase, and (3) a post-treatment, follow-up phase. During the baseline phase, the researcher asks the participant to complete daily dependent measures for anywhere from one to four weeks. For example, in a study of depression treatment, the participant might complete four weeks of daily mood ratings. These ratings allow the researcher to establish a "stable" baseline—that is, a picture of the degree to which the participant's mood varies during a typical four-week period. The participant continues to complete these daily measures once treatment has begun. A follow-up data collection phase, during which the participant resumes the daily dependent measures, occurs several weeks or months after treatment has ended. This phase allows the researcher to observe whether the participant maintained any gains she may have made during the treatment phase.

The large number of observations used in time-series designs allows researchers to determine whether changes in dependent variables over time are statistically significant. In order to do this, the statistical test used must account for autocorrelation. According to Borckardt et al. (2008) "a series of observations...is said to be autocorrelated if the value of one observation depends (at least in part) on the value of

one or more of the immediately preceding observations. Later observations are explained by earlier ones” (p. 82). Statistical programs such as Simulation Modeling Analysis (SMA; Borckardt et al., 2008) and Interrupted Time Series Analysis (ITSACORR; Crosbie, 1993) can be used to compare the slope and level of the dependent variable before and after the intervention, while controlling for the effects of autocorrelation.

Goals of the current study

The current study investigates the therapeutic value of TA when it is used with couples. Specifically, this study explores whether a TA intervention, conducted by psychologists at the Center for Therapeutic Assessment, helps couples and therapists resolve therapeutic impasses, leading to greater satisfaction with both therapy and the romantic relationship. Although case studies suggest TA may be a powerful intervention for couples, empirical research is needed. The current study is the first to systematically examine whether TA does, in fact, help couples make gains in couples therapy.

The four main hypotheses being investigated in the current study are highlighted in bold. Finn (2007) noted that couples who participate in TA tend to report feeling happier and more satisfied in their relationships as a result of the intervention. It was therefore hypothesized that **the TA intervention would lead to increases in reported relationship satisfaction and decreases in reported disaffection and interpersonal conflict**. In addition, results from preliminary studies have indicated TA may lead to greater psychological wellbeing (Finn & Tonsager, 1992; Newman & Greenway, 1997). Therefore, a second hypothesis was that **participants would report decreases in the frequency and severity of symptoms of psychological problems, such as depression and anxiety**. Finn (2007) has also observed that TA appears to help clients make more satisfactory progress in therapy. Based on these observations, a third hypothesis of the current study was that, after the TA intervention, **both the couples and their therapists would report an increase in their satisfaction with the progress being made in couples therapy**. Finally, given that previous research has indicated most clients tend to be highly satisfied with their experiences participating in TA (Finn, 2007), and that

satisfaction with TA predicts positive therapeutic outcomes (Finn & Tonsager, 1992), it was hypothesized that **participants in the current study would report high levels of satisfaction with the TA intervention**. Time-series analysis was used to test whether there was a significant change in daily measures of relationship satisfaction between the baseline, intervention, and follow-up phases of the study (Borckardt, Nash, Murphy, Moor, Shaw, & O'Neil, 2008). Visual analysis was also used to examine whether couples reported changes on global measures of relationship satisfaction, psychological symptoms, and therapy progress (Kazdin, 1983).

Chapter III: Methodology

Participants

Participants were three married or cohabitating heterosexual couples who, at the time of the study, were already involved in ongoing couples therapy. Therapists in the community were asked to identify couple clients who were currently “stuck” in therapy. Specifically, therapists were asked to identify cases in which clients had attended at least 12 therapy sessions and in which both the clients and the therapist had acknowledged that progress was slow or unsatisfactory, or in which the therapist thought an assessment consultation might be helpful for the therapist and clients.

Participants experiencing symptoms of psychosis were excluded from the study. In addition, couples who were not currently living together (e.g., going through a trial separation) or who had been living together for less than one year were excluded. Couples in which one or both partners frequently left town for more than two days each month were also be excluded. Finally, Finn (personal communication, August 5, 2008) has observed that some couples participate in TA because they are seriously contemplating divorce/separation and want help in making this decision. In order to examine changes in relationship satisfaction over time, the current study targeted couples in which both partners had expressed a commitment to working on improving the relationship. Although couples who are contemplating separation might potentially benefit from TA, they were excluded from the current study.

For their participation in the first two phases of the study, participants were given the Therapeutic Assessment intervention at no charge. Couples were also given a \$20 gift card after completing the follow-up phase of the study. In addition, therapists were be given a \$20 gift card for their participation in the study.

Data collection procedure

The data collection procedure for the current study is summarized in Figure 3.1. For the initial, pre-baseline meeting with the couples, the researcher met with couples in a laboratory at the University of Texas at Austin. Participants then completed brief, daily

measures of relationship satisfaction throughout three phases of the study: a pre-intervention baseline phase (duration: four weeks), an intervention phase (duration: six to seven weeks), and a post-intervention follow-up phase, which will begin immediately after the couples' final TA session (duration: four weeks). In addition, the couples and their therapists completed longer, standardized measures of relationship satisfaction, psychological symptoms, and therapy progress at three time points: pre-intervention, post-intervention, and at the end of the follow-up phase (four weeks after the last TA session).

Figure 3.1 *Data collection procedure.*

<i>Research phase</i>	<i>Description</i>	<i>Duration</i>
Pre-baseline	<ul style="list-style-type: none"> • Researcher went over informed consent with couple. • Researcher interviewed couple about presenting problems. • Each member of the couple completed global measures of relationship functioning, self-esteem, and psychological symptoms (DAS, MDS, and BSI). • Based on interview and DAS and MDS scores, the researcher developed brief measures for each member of the couple to complete on a daily basis. 	1 day
Baseline	<ul style="list-style-type: none"> • Each member of the couple completed brief daily measures of relationship functioning. • Each member of the couple, as well as the couple's therapist completed a weekly measure of therapy satisfaction (SEQ). 	4 weeks
Intervention	<ul style="list-style-type: none"> • Couple attended TA sessions (one session per week). • Couple continued to complete daily measures of relationship functioning. • Couple and therapist continued to complete the SEQ on a weekly basis. • At the end of the intervention phase, the couple once again completed the DAS, MDS, and BSI. Couple also completed the AQ-2 and an open-ended feedback questionnaire. 	6-7 weeks
Follow-up	<ul style="list-style-type: none"> • Couple completed daily measures of relationship functioning. • Couple and therapist completed SEQ on a weekly basis. • At the end of the follow-up phase, the couple once again completed the DAS, MDS, & BSI. 	4 weeks

Note. AQ-2=Assessment Questionnaire-2; BSI=Brief Symptom Inventory; DAS=Dyadic Adjustment Scale; MDS=Marital Disaffection Scale; SEQ=Session Evaluation Questionnaire

Description of the TA intervention

In the current study, couples participated in a brief TA intervention. The intervention comprised five sessions, which took place over the course of 6-7 weeks. The intervention followed Finn's (2007) guidelines for conducting TA with couples; however, it was relatively brief, compared to the TAs normally conducted at the Center for Therapeutic Assessment. In a longer, more typical TA, assessors would administer individual Rorschachs to each partner, and also select additional individual- and couple-level assessment tools to administer based on the couple's presenting concerns and assessment questions.

The assessors were three psychologists from the Center for Therapeutic Assessment (CTA) in Austin, Texas. The first assessor was a licensed psychologist with over 15 years of experience in the research and practice of Therapeutic Assessment; he is also the founder of the CTA. The second assessor was a licensed early career psychologist. The third assessor was a psychologist who, at the time of the study, was completing her postdoctoral training and was supervised by the first assessor. All three assessors regularly met with one another, as well as other members of the CTA, for case consultation. Each assessor was assigned to work with one couple. Each TA session was videotaped. A week-by-week description of the intervention is described in Figure 3.2.

Figure 3.2. *Overview of Therapeutic Assessment intervention.*

<i>Week</i>	<i>Intervention</i>
1	Couple met with assessor for initial, hour-long session. Each member of the couple was given the MOS form to complete at home. Couple's therapist met with assessor for a 30-60 minute session.
2	First member of the couple met with assessor one-on-one for an hour-long session.
3	Second member of the couple met with assessor one-on-one for an hour-long session. Both members of the couple came to the CTA to complete the MMPI-2, and FAM-III. Each member of the couple returned the MOS form to the assessor.
4	Couple took the Consensus Rorschach in a session lasting 2-2½ hours.
5	Couple's therapist met with the assessor one-on-one to go over findings from the assessment. Couple then met with the assessor for a two-hour-long summary/discussion session. The couple's therapist joined the couple for this session.
6 or 7	Couple received feedback letter in the mail.

Assessment instruments used in the TA intervention

The four assessment instruments that were used in the TA intervention were the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom, & Kaemmer, 2001), the Family Assessment Measure-III (FAM-III; Skinner, Steinhauer, & Santa-Barbara, 1983; Skinner, Steinhauer, & Sitarenios, 2000), the Memories of Spouse Procedure (MOS; Bruhn, 1992a), and the Rorschach Inkblot Technique (Exner, 1993). Here, I will provide a brief overview of each of these instruments.

Minnesota Multiphasic Personality Inventory-2 (MMPI-2)

The MMPI-2 is the most widely-used clinical testing/personality instrument in psychological assessment (Butcher & Williams, 2000). It consists of 567 true-false items and comprises 3 validity scales, 10 basic scales, 18 supplementary scales, and 15 content scales. The ten basic scales are those most commonly-used in TA (Finn 1996, 2007). The MMPI-2 was normed on a randomly chosen sample of 2,600 adults from various regions

of the United States and its reliability and validity have been supported in numerous studies over the years (see Groth-Marnat, 2003, for an overview). When using the MMPI-2 in TA, the assessor uses the results of the test to form hypotheses about the client's test-taking attitude, level of distress and disturbance, major symptoms, underlying personality, behavior in relationships, implications for treatment, and diagnostic impressions (Finn, 1996).

Family Assessment Measure—III (FAM-III)

The FAM-III (Skinner et al., 1983, 2000) is a 134-item self-report measure based on Steinhauer, Santa-Barbara, and Skinner's (1984) Process Model of Family Functioning, which describes family functioning as occurring along seven dimensions: task accomplishment, role performance, communication, affective expression, involvement, control, and values and norms. The FAM-III has three scales: 1) whole family system (General scale; 50 items), 2) dyadic relationships (Dyadic scale; 42 items), and 3) individual functioning (Self-Rating scale; 42 items). Each of the three scales assesses family functioning on all seven dimensions. Raw scores are converted to T-scores, with scores above 60 indicating problematic family functioning. Assessors can choose to use one or all of the scales as they see fit. In the current study, only the Dyadic and Self-Rating scales were used.

Skinner et al. (2000) noted that the FAM-III "can generate an unusually rich picture of a couple's relationship if, in addition to the partners using the dyadic scale, they also complete...self scales, which demonstrate how they see—or don't see—their part in the couple's problem" (p. 197). In their 1983 study, Skinner et al. found that scores the FAM-III successfully discriminated between "problem" families (i.e., families in which one or more members were receiving professional help for psychiatric or emotional problems) and "non-problem" families. Skinner et al. (1983) also reported high internal consistency reliability for scores on both the Dyadic scale (Cronbach's $\alpha=.95$) and the Self-Rating scale (Cronbach's $\alpha=.89$).

Memories of Spouse Procedure (MOS)

The Memories of Spouse Procedure (MOS; Bruhn, 1992a) is based on Bruhn's (1992b, 1992c) Early Memories Procedure (EMP), a projective, autobiographical writing task. The EMP asks respondents to write about a number of early childhood memories (e.g., first five earliest memories, earliest memories of parents, earliest memory of going to school). Interpretation of these memories uses Bruhn's cognitive-perceptual theory, which posits that "we recall incidents that mirror and substantiate our current perceptions and beliefs and 'overlook' or 'forget' incidents that are inconsistent with present beliefs" (Bruhn, 1992b, p. 13). That is, the memories that we tend to recall, as well as the way in which we talk or write about these memories, says a great deal about the way in which we view ourselves and the world. When reading through clients' early memories, assessors examine the content of the memories to form hypotheses about clients' approaches to problem-solving, unresolved emotional and interpersonal issues, and perceptions of self and others.

In the current study, participants were asked to write about the six memories specific to their relationship with their current partner or spouse: (1) "A Time You Felt Close to Your Spouse," (2) "A Time You Felt Cared For by Your Spouse," (3) "A Time You Felt Protected by Your Spouse," (4) "A Time You Felt Hurt by Your Spouse," (5) "A Time You Felt Betrayed by Your Spouse," and (6) "A Time You Felt Unprotected by Your Spouse." These prompts, which are those currently being used by psychologists at the Center for Therapeutic Assessment, were an adaptation of Bruhn's original (1992a) prompts. After writing about each memory, participants were asked to respond to four questions: (1) What is the clearest part of the memory? (2) What is the strongest feeling in the memory? (3) If you could change the memory in any way, what would that be? and (4) What was your approximate age at the time of the memory? Assessors using the MOS procedure can use clients' written memories to form hypotheses about the dynamics at work in the relationship. In the current study, assessors referred to these memories in the Consensus Rorschach session and/or the summary and discussion session.

Rorschach Inkblot Technique

The Rorschach (Exner, 1993) is currently one of the most widely-used psychological assessment techniques (Groth-Marnat, 2003). First published by Hermann Rorschach in 1921, the Rorschach test consists of ten symmetrical inkblots that are presented to a client with the prompt, “What might this be?” The administration consists of both a “free association” phase, in which clients give their free responses to the cards, and an inquiry phase, during which the assessor asks questions about the determinants and locations of the responses. Before the 1970s, there was no standardized method for administering or interpreting the Rorschach. In 1974, however, John Exner, after collecting normative data on the Rorschach, published his Comprehensive System, a set of detailed procedures for scoring and interpreting clients’ responses. Over the past three decades normative data have been revised several times, such that a large body of empirical evidence now supports the utility of the Rorschach in identifying psychological problems such as depression and psychosis, as well as certain personality characteristics (e.g., narcissism, paranoia) and interpersonal styles (e.g., passivity, aggression). The guidelines for administering, scoring, and interpreting the Rorschach are quite complex, and it is recommended that students undergo 30-50 hours of class instruction and at least one year of supervision before they are proficient in its use (Nurse, 1999). When clinicians have undergone proper training, inter-rater reliability tends to be high. In Meyer and colleagues’ (2002) review of eight samples of Rorschach data, median intraclass correlation values ranged from .82 to .97. Other studies have supported the validity of the Rorschach. In a meta analysis, Hiller, Rosenthal, Bornstein, Berry, and Brunell-Neuleib (1999) found that, on average, the Rorschach tends to have good criterion-related validity ($r=.29$), and that this value is comparable to that of the MMPI (average $r=.30$). In addition, Ganellen’s (1996) meta-analysis found that the Rorschach tends to have higher specificity than the MMPI and the Millon Clinical Multiaxial Inventory-II (MCMI-II) in predicting diagnoses of depression and higher specificity *and* sensitivity than these two measures in predicting diagnoses of psychosis.

In the current study, assessors used the Consensus Rorschach procedure (see Chapter II for a description). During this procedure, assessors focused primarily on observations of process (i.e., how the couple worked together, reached agreements, and solved problems); however, assessors made note of the content, location, form and developmental quality of the responses. Although the assessors did not formally score couples' Rorschach protocols, they made use of Exner's scoring system to inform their hypotheses about each partner's psychology. For example, morbid or shading responses may indicate underlying depression in one or both of the partners, while responses of poor form quality might suggest impaired reality testing.

Integrity checks

Treatment integrity checks are often performed in outcome research to assess the degree to which an intervention was performed as intended. In the current study, to ensure that assessors conducted the assessments according to the principles of Therapeutic Assessment, as described by Finn (2007), an advanced doctoral student who was trained and experienced in the use of Therapeutic Assessment viewed segments of videotapes of the sessions. Using a rating form designed for this study (see Appendix I), the rater evaluated both treatment adherence (i.e., Did the assessors follow Finn's steps for conducting TA?) and assessor competence (i.e., How skillful were the assessors in their use of TA techniques?) (Heppner, Kivlighan, & Wampold, 1999). To ensure that the assessors were aware of the criteria on which they were to be rated, each assessor was given copies of the integrity rating forms before conducting the TA intervention.

Measures

Demographic Questionnaire. A brief demographic questionnaire developed for this study (Appendix A) was used to collect data on participants' age, race/ethnicity, socioeconomic status, and length of their current relationship.

Couples' daily measures

Each member of the couple completed brief daily measures of relationship functioning. Participants were asked to complete these measures at approximately the same time every day (e.g., after dinner or before bedtime). Although couples were given the option of completing either pen and paper measures or online measures using a secure website, all three couples opted to complete the measures online. These daily measures contained items that were chosen based on the couples' initial scores on standardized measures of relationship functioning (described below), as well as an initial interview with the researcher. In order to tailor the measures to each couple's presenting problem, most of these items varied from participant to participant and were developed based on the Time 1 interview with the couple, incorporating each participant's own language whenever possible. For example, a participant whose presenting concerns included feeling a great deal of tension between herself and her husband was asked to respond to the following item each day: "Please rate the degree of tension between you and Ray today." (1=no tension at all; 7=extreme tension). Participants were shown a draft of these measures before the daily data collection phase began. If an item did not meet a participant's approval (i.e., the participant does not believe the item accurately represents one of the target problems), the researcher and the participant collaborated in revising the item.

In addition to answering individually-tailored questions like the ones listed above, all couples were also asked to respond to the following three questions on a daily basis:

- Please rate your general feelings toward your partner today? (on a scale of 1-7; 1 = extremely positive, 7 = extremely negative)
- How close did you feel to your partner today? (on a scale of 1-7; 1 = extremely close, 7 = extremely distant)
- How hopeful did you feel about your relationship today? (on a scale of 1-5; 1 = very hopeful, 5 = not at all hopeful)

Finally, the daily questionnaire contained an open-ended comments section where participants could choose, if they wished, to write about anything that happened that day that may have affected their relationship satisfaction ratings.

It should be noted that because these daily questions were developed specifically for this study, and most questions were tailored to each participant's presenting problems, the reliability and validity of the measures could not be established. It will be important to keep this in mind when interpreting the results from the daily measures.

Standardized measures of relationship functioning and emotional wellbeing

The following are longer measures that were collected at the beginning of the study (Time 1) and again at the end of the intervention (Time 2) and follow-up (Time 3) phases.

Dyadic Adjustment Scale (DAS; Spanier, 1976; Appendix B). The DAS, one of the most widely-used self-report measures of relationship adjustment in couples, comprises 32 Likert-type items and four factors. The first factor, Dyadic Consensus, measures the degree to which couples agree or disagree with one another on various issues (e.g., handling family finances or household tasks). Items loading on the second factor, Dyadic Satisfaction, ask respondents to rate their level of happiness or satisfaction in the relationship (e.g., "In general, how often do you think that things between you and your partner are going well?"). The third factor, Dyadic Cohesion, contains items on which respondents rate the frequency with which they engage in pleasant or productive activities with their partners (e.g., "Laugh together," or "Work together on a project"). The fourth factor, Affectional Expression, assesses respondents' satisfaction with the quality and frequency of demonstrations of affection in the relationship (e.g., "being too tired for sex" or "not showing love"). Higher scores on the DAS indicate more positive dyadic adjustment in a couple. In most studies using the DAS, total scores below 92 are thought to indicate relationship distress, while scores above 107 are thought to indicate the absence of significant distress (Sabourin, Valois, & Lussier, 2005). Spanier (1976)

found that each item on the DAS successfully discriminated between divorced and married respondents. Spanier also found scores on the DAS were significantly correlated with scores on the Locke-Wallace Marital Adjustment Test (Locke & Wallace, 1959), another commonly-used measure of marital adjustment ($r=.86$ among married respondents and $.88$ among divorced respondents). In a reliability generalization meta-analysis of the DAS, Graham, Liu, and Jeziorski (2006) found that the mean overall Cronbach's alpha reported across studies for the DAS was $.915$. Reliability for scores on individual subtests was also generally good (mean Cronbach's alpha across studies: Dyadic Consensus= $.872$, Dyadic Satisfaction= $.848$, Dyadic Cohesion= $.789$, Affectional Expression= $.714$).

Marital Disaffection Scale (MDS; Kayser, 1996; Appendix C). The MDS measures "disaffection," or the degree to which couples feel they have lost positive emotions for one another (e.g., love, caring, affection, closeness) over time. The scale contains 21 Likert-type items on which respondents indicate the degree to which they agree or disagree with statements such as "Apathy and indifference best describe my feelings toward my spouse," or "My love for my spouse has increased more and more over time." This unidimensional scale is scored such that higher scores indicate higher levels of disaffection. Scores range from 21-84, and Kayser (1996) assigned the following classifications to the following score ranges: 21-26=low disaffection; 27-34=below average disaffection; 35-42=average disaffection; 43-54=above average disaffection; and 55-84=high disaffection. Kayser (1996) found that the MDS successfully discriminated between couples who were in marital counseling and those who were not in counseling. In Kayser's sample, scores on the MDS were also highly correlated with scores on another measure of disaffection, and were negatively correlated with respondents' reports of marital happiness and closeness. Studies using the MDS have reported high internal consistency reliability, with Cronbach's alphas ranging from $.87$ to $.97$ (Kayser, 1996; Priest & Thein, 2003).

Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983; Appendix D). The BSI, a 53-item self-report measure of psychological symptomatology, is a short version of the Symptom Checklist-90-Revised (SCL-90-R; Derogatis, Rickels, & Rock, 1976). Respondents are presented with a list of problems (e.g., “feeling lonely” or “nervousness or shakiness inside”) and are asked to use a 5-point Likert scale to rate the degree to which each problem bothered them in the last week, with responses ranging from “not at all” to “extremely.” The BSI contains nine primary symptom dimensions: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. In addition, the BSI contains three global indices of distress: a General Severity Index (GSI), a Positive Symptoms Distress Index (PSDI), and a Positive Symptom Total (PST). According to Derogatis and Melisaratos, “the GSI is the single best indicator of current distress levels, and should be utilized in most instances where a single summary measure is required” (p. 597). On each index, higher scores indicate higher levels of symptomatic distress. Derogatis (1993) developed norms for adult nonpatients, psychiatric outpatients, and psychiatric inpatients. These norms can be used to convert raw scores to T scores.

Results from Derogatis and Melisaratos’ (1983) study support the reliability and validity of scores on the BSI. In their sample of 1002 outpatients, the authors reported Cronbach’s alphas ranging from .71 to .85 on scores on each of the nine dimensions. In a sample of 60 non-patient respondents, test-retest reliability over a two-week period ranged from .68 (for scores on the somatization scale) to .91 (for scores on the phobic anxiety scale). In addition, the authors found that in a sample of 209 symptomatic volunteers, scales on the BSI were correlated with similar scales on the MMPI.

Satisfaction with the Therapeutic Assessment intervention

Assessment Questionnaire-2 (AQ-2; Finn, Schroeder, & Tonsager, 1994; Appendix E). The AQ-2 is a standardized, 48-item instrument designed to measure clients’ satisfaction with psychological assessment. Respondents rate their satisfaction on four subscales: a) how much they learned about themselves from the assessment (new

self-awareness/understanding), b) how well understood they felt after the assessment (positive accurate mirroring), c) how positively they experienced their relationship with the assessor (positive relationship with assessor), and d) how negatively they experienced the assessment (negative feelings about the assessment). An overall, general satisfaction (GS) score is also calculated. The AQ-2 has demonstrated adequate internal consistency reliability, with Cronbach's alphas ranging from .76-.93 for scores on each of the four subscales (Finn et al., 1994; Newman & Greenway, 1997). Finn et al. (1994) also reported good test-retest reliability (range: .75-.84) in their sample of college students, outpatient clients, and psychiatric inpatients.

Feedback Questionnaire—Client Form (Appendix F). The Feedback Questionnaire—Client Form contains eight open-ended questions asking clients to give written feedback about their experience during an assessment. Six items from this questionnaire are routinely used by psychologists at the Center for Therapeutic Assessment (Finn, personal communication, August 5, 2008). The questionnaire contains questions such as “How well did the assessment meet your expectations?” and “What suggestions do you have for the way we do assessments?” For the purposes of this study, two items pertaining to the participants' experience in *couples'* TA were added to the questionnaire: “How do you think the assessment has affected your relationship with your partner?” and “How do you think the assessment has affected your experience in couples therapy?”

Feedback Questionnaire—Therapist Form (Appendix G). The Feedback Questionnaire—Therapist Form contains six open-ended questions asking referring therapists to give written feedback about their experience of the TA. This questionnaire is routinely used by psychologists at the Center for Therapeutic Assessment (Finn, personal communication, August 5, 2008). The questionnaire contains questions such as “What part(s) of the assessment was most useful to you and your clients?” and “What part was least useful?”

Satisfaction with weekly couples therapy

Session Evaluation Questionnaire—Client Form (SEQ-C; Stiles, Reynolds, Hardy, Rees, Barkham, & Shapiro, 1994; Appendix H). The SEQ-C was designed to assess clients' experience during a single psychotherapy session and contains 21 bipolar adjective scales which are rated on a scale of 1-7. The measure is divided into two sections: session evaluation and post-session mood. The first section (session evaluation) contains two subscales: 1) Depth, which measures the degree to which the client experienced the session as powerful/valuable versus weak/worthless; and 2) Smoothness, which measures the degree to which the client experienced the session as smooth/comfortable versus rough/uncomfortable. The second section (post-session mood) contains two subscales: 1) Positivity, on which clients rate the degree to which they feel happy/confident/pleased versus sad/afraid/angry; and 2) Arousal, on which clients rate the degree to which they feel aroused/energetic/excited versus quiet/peaceful/calm. The SEQ-C also contains one global impression item on which clients are asked to rate the degree to which they experienced the session as "good" or "bad." Stiles et al. (1994) reported high internal consistency reliability for scores on all four subscales of the SEQ (Depth=.90, Smoothness=.92, Positivity=.90, Arousal=.80). Convergent validity was also found to be good, with scores on the SEQ significantly correlating with scores on other measures of psychotherapy session impact. Although the SEQ was originally developed for use with individual therapy clients, it has since been used in a variety of settings, including group, family, and marital therapy, and can be completed by both clients and therapists (Stiles, Gordon, & Lani, 2002).

Session Evaluation Questionnaire—Therapist Form (SEQ-T; Stiles, Reynolds, Hardy, Rees, Barkham, & Shapiro, 1994). In the current study, the SEQ-T was identical to the first section of the SEQ-C (session evaluation), with the second section (post-session mood) omitted. Each couple's therapist was asked to complete the SEQ-T immediately after each therapy session.

Hypotheses and Analyses

The current study used case-based, time-series analysis (Borckardt, Nash, Murphy, Moor, Shaw, & O'Neil, 2008) to examine whether each member of each couple experienced an increase in relationship satisfaction as a result of the Therapeutic Assessment intervention. In addition, visual analysis (Kazdin, 1982) was used to examine whether participants reported changes in relationship functioning, psychological symptoms, and satisfaction with couples therapy.

Hypotheses

Hypothesis 1

- A. After participating in the TA intervention, participants will report improvements on ideographic (individualized), daily measures of relationship satisfaction (measures to be determined for each participant, depending on the presenting concern).
- B. These improvements in relationship satisfaction will be sustained, or will continue to increase at follow-up.

Hypothesis 1 Data Analysis

- A. Simulation modeling analysis (SMA) was used to determine whether there was a change in the level of plotted ratings of relationship satisfaction between the baseline and intervention phases. SMA was also used to compare ratings at baseline with ratings at follow-up. SMA was chosen because it allows researchers to conduct time series analyses with relatively few observations per phase (Borckardt et al., 2008).

Hypothesis 2

- A. After participating in the TA intervention, participants will report increases in relationship satisfaction as measured by nomothetic (global) measures of relationship functioning.
- B. These increases will be sustained or will continue to improve at follow-up.

Hypothesis 2 Data Analysis

- A. Visual analysis was used to compare DAS and MDS scores obtained at Time 1 to those obtained immediately after the TA intervention and to those obtained at the end of the follow-up phase.
- B. Scores on both measures were examined to determine whether participants moved from the distressed range (<92 on the DAS; >42 on the MDS) to the non-distressed range between baseline and follow-up.

Hypothesis 3

- A. Participants will report improved ratings of progress in couples therapy.
- B. Couples' therapists will also report improved ratings of their clients' progress in therapy.

Hypothesis 3 Data Analysis

Visual analysis was used to compare both SEQ-C and SEQ-T scores obtained at Time 1 to those obtained immediately after the TA intervention and to those obtained at the end of the follow-up phase.

Hypothesis 4

- A. Participants will report decreases in psychological symptoms after participating in the TA intervention.
- B. These decreases in symptomatology will be sustained, or will continue to decrease at follow-up.

Hypothesis 4 Data Analysis

- A. Visual analysis was used to compare BSI scores obtained at Time 1 to those obtained immediately after the TA intervention and to those obtained at the end of the follow-up phase.
- B. Scores on the BSI were examined to determine whether participants moved from the distressed range (T score >65) to the non-distressed range between baseline and follow-up.

Hypothesis 5

Participants will report feeling satisfied with the TA intervention. Specifically, participants will report above average levels of satisfaction on all four subscales of the AQ-2 (T-scores above 50). In addition, it is hypothesized that, overall, participants' written feedback on the Feedback Questionnaire will be positive.

Hypothesis 5 Data Analysis

Scores on each of the four subscales of the AQ-2 were examined to determine whether participants' T-scores fell above or below 50. In addition, the researcher conducted a content analysis of the participants' written responses to the Feedback Questionnaire.

Hypothesis 6

Couples' therapists will report feeling satisfied with the TA intervention. Specifically, therapists will report feeling that the intervention aided the couples' therapy.

Hypothesis 6 Data Analysis

The researcher conducted a content analysis of the therapists' written responses to the therapist version of the Feedback Questionnaire.

Chapter IV: Results and Discussion

Overview of results to be presented

In this chapter, results and discussion will be presented for each couple separately, essentially treating each couple as its own case study. The results section for each couple will begin with a description of the couple and their presenting problems. The treatment integrity ratings for each Therapeutic Assessment will also be provided. Qualitative information obtained as part of the couples' daily surveys will then be summarized. Next, results from the time-series phase effect analyses (described in more detail in the following section) will be presented, followed by the results from the measures of global relationship satisfaction (DAS and MDS), psychological symptoms (BSI), and progress in couples therapy (SEQ-C). Couples' quantitative (AQ-2 scores) and qualitative ratings of their satisfaction with the TA will then be presented. Finally, the couples' therapists' quantitative ratings of their clients' progress in couples therapy (SEQ-T) and their qualitative feedback about the TA will be summarized and briefly discussed.

Time-series phase effect analyses of daily measures

For each participant's daily measures, simulation modeling analysis (SMA) was used to assess whether the level and slope of plotted ratings of relationship satisfaction changed significantly between the baseline and TA phases, and between the baseline and follow-up phases. In the following sections, the results of the phase-effect analyses will be presented for each question that each participant answered on a daily basis. In each case, I will indicate whether there was a significant change in both the *level* and the *slope* of the dependent variable across phases. When assessing for these changes, three statistics will be presented: (1) r (Pearson's correlation between the dependent variable and the phase vector), (2) $r(\text{Lag } 1)$ (autocorrelation statistic), and (3) the p-value, which is the estimate of the probability of the observed phase-effect occurring by chance. The mean, standard deviation, and ordinary least squares slope for ratings on each daily

question (within each phase of the study and across phases) presented in table form in Appendix J.

In the current study, the Expectation-Maximization Algorithm (Dempster, Laird, & Rubin, 1977) was used to estimate missing values for each of the daily measures. Velicer and Colby (2005) found that the Expectation-Maximization (EM) procedure provided the most accurate estimate of missing time-series data values when compared to other methods, such as deletion, mean substitution, and mean of adjacent observations. In addition, these authors found that the EM procedure provided accurate missing data estimates for data sets in which up to 40% of the data were missing completely at random. In the current study, Little's MCAR test (Little, 1988) was used to test the assumption that data were missing completely at random for each participant. P-values of greater than .05 indicated that the MCAR assumption was met.

Sample characteristics

Three couples and their therapists participated in the study. Two of the couples were married and one was cohabitating. The average age of the couple participants was 43.67 and the standard deviation was 13.9. Two of the six participants had graduate degrees and the other four had Bachelor's degrees. Five of the participants identified as European American/White and one identified as Latina. The number of years together as a couple ranged from two to fifteen years. Each couple will be discussed in further detail in the following sections.

The therapist participants were all experienced psychotherapists working in private practice. All three were women. One therapist was a doctoral level psychologist and the other two were clinical social workers. All three therapists were well-regarded in the community as being "expert" couples therapists; in particular, the therapist for Couple #2 (George and Susan) was well-known for her writing and clinical expertise in the area of couples therapy.

All names and potentially identifying information, for both couples and their therapists, have been changed to disguise the participants' identity.

Couple #1: Mary and Ray

Mary and Ray were in their late 30s/early 40s, had been married for over ten years, and had two children, ages 8 and 11, at the time of the study. Ray, a White man with a Bachelor's degree, owned a small business. Mary, a Latina woman, also had a Bachelor's degree and was a stay-at-home mother. At the outset of the study, they had been seeing their couples therapist, Jane, for approximately nine months, and had originally presented for therapy due to what they called a "communication breakdown." They reported they had been experiencing a great deal of tension and conflict with one another, particularly around Ray's friendship with another woman, of whom Mary had become quite jealous. Mary felt as though Ray was no longer open and honest with her, and felt their relationship had come to be characterized by a "lack of trust." She was also concerned about Ray's drinking, which had increased somewhat over the past year. She said her problems with Ray caused her to question her self-worth as a mother and left her feeling depressed. Ray felt that Mary could be overbearing and needy, and he found himself wanting to spend more time away from the home, with friends. He reported he had grown tired of the frequent arguments with Mary, and had increasingly begun to keep his thoughts to himself for fear that speaking would lead to yet another argument. Both Mary and Ray expressed concern that their arguments were having a negative impact on their children.

Mary and Ray reported couples therapy had helped them improve communication with one another and caused their fights to be somewhat less explosive. At their Time 1 interview, however, they reported they still felt as though they were "spinning [their] wheels" and "not getting to the root of the problem."

After the Time 1 interviews with Ray and Mary, the researcher used information from the interview, as well as the couple's responses to the DAS, MDS, and BSI, to develop a set of questions to which each partner would respond on a daily basis throughout the study. To ensure that the questions accurately represented the couple's presenting problems, the researcher sent drafts of the questions to both Mary and Ray and

invited them to edit or amend the questions as they saw fit. Each of these questions is presented along with the time-series results.

Integrity rating of TA sessions

Treatment integrity ratings indicated Mary and Ray’s assessor conducted the assessment according to the principles of Therapeutic Assessment as described by Finn (2007) (see Table 4.1).

Table 4.1 *Treatment integrity ratings—Mary and Ray*

Initial session with couple	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. The assessor sought assessment questions from both clients.					X
2. The assessor asked the clients about their “best guess” answers to at least one of these questions.					X
3. The assessor let the clients know that he/she would be asking for their input/collaboration throughout the assessment process.				X	
4. The assessor talked to the clients about the tests they would be taking.					X
5. The assessor demonstrated empathy.					X
6. The assessor demonstrated good listening skills.					X
7. The assessor encouraged the clients to ask questions about the assessment.					X
8. If a client asks a non-systemic question about his/her partner, the assessor helped to rephrase this as a systemic question.					n/a
Consensus Rorschach/Assessment Intervention session	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. During the Consensus Rorschach task, the assessor asked the couple to come up with responses that they could both see and agree on.					X

Table 4.1, cont.

2. After administering the Consensus Rorschach, the assessor asked the clients to talk about their experience during the task (e.g., thoughts, feelings, observations).					X
3. The assessor attempted to make connections between what happened during the Consensus Rorschach task and what happens in other areas of the couple's life together.					X
4. The assessor asked the couple to try out a new approach to the Consensus Rorschach (using a card that was not used during the first administration).					X
5. The assessor demonstrated empathy.					X
6. The assessor demonstrated good listening skills.					X
7. The assessor made systemic interpretations about how each person contributes to the patterns revealed.					X
Summary/Discussion session	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. The assessor appeared to begin by presenting Level 1 findings, followed by Level 2 and (if appropriate) Level 3 findings.					X
2. The assessor encouraged the clients to revise or amend the findings.					X
3. The assessor demonstrated empathy.					X
4. The assessor demonstrated good listening skills.					X
5. The assessor seemed to give equal emphasis to each partner's contribution to the relationship struggles.					X
6. The assessor asked clients to give real life examples of assessment findings that were discussed.					X

Results: Mary and Ray

Qualitative information from daily surveys

The following prompt was included on each of Mary and Ray's daily surveys: "If anything happened today that you think affected your feelings about your relationship in any way, feel free to mention it in the box below." To provide a sense of some of the situational factors that may have affected or been affected by the couple's experience of the TA intervention, their qualitative responses to this prompt will be summarized here.

Throughout the baseline phase of the study, both Mary and Ray reported arguments and tension with one another in the comments section of their daily surveys. For example, on the first day of the baseline phase, Mary noted, "I believe that Ray left this morning without saying goodbye or kissing me goodbye. He says he did and that I just forgot. We argued a little about this. He also told me that he is leaving town for 3 days with his business partner. This makes me feel tense and uncomfortable." Similarly, one week into the baseline phase, Ray noted, "Always tension about having a female business partner. This evening Mary annoyed about a business trip to [City] for just the day, tomorrow." During the second week of the baseline phase, Mary reported, "Our therapy session did not go very well and we argued afterward. The evening was a bit tense too. We're not really talking tonight. We're also both very tired, especially him." Throughout this phase, both partners also noted some positive experiences with one another, such as an enjoyable family trip and a fun date night together.

On the first day of the TA phase, Mary's written comment indicated she had a very positive response to their first meeting with the TA assessor:

It was a bit of an eye-opening experience for me today. We met with Dr. F [assessor for the TA] and issues were discussed that helped me understand Ray better. Although, the discussion was emotional, somewhat painful, and humbling, it also made me realize how much I love Ray and how much I want our marriage to last forever.

Four days later, Mary wrote the following: “I discovered Ray is having an affair. I’m trying to function.” That same day, Ray wrote, “Extremely destructive incident by me. Bad and explosive fight ensued all day Sunday. Very serious and wrenching.” After this, Ray did not provide any more qualitative comments for the remainder of the study. A few days later, Mary reported, “We’re talking more, and I’m feeling a little stronger.” Her comments throughout the rest of the TA phase indicated she experienced many ups and downs in her relationship with Ray. At one point, for example, she expressed a fear that couples therapy was not working and reported feeling “low and confused.” Later in the phase, she reported, “Ray wrote something really special and beautiful in my Mother’s Day card. Everything he did today meant so much to me.”

On the first day of the follow-up phase, Mary reported she and Ray argued in the evening, but apologized to one another afterwards. Approximately one week later, she reported that they “argued a lot today” and stated, “I feel he is very insincere and only shows affection and caring when it’s convenient for him.” She did not write any more comments for the remaining three weeks of the follow-up phase.

Time-Series Phase Effect Analyses—Ray

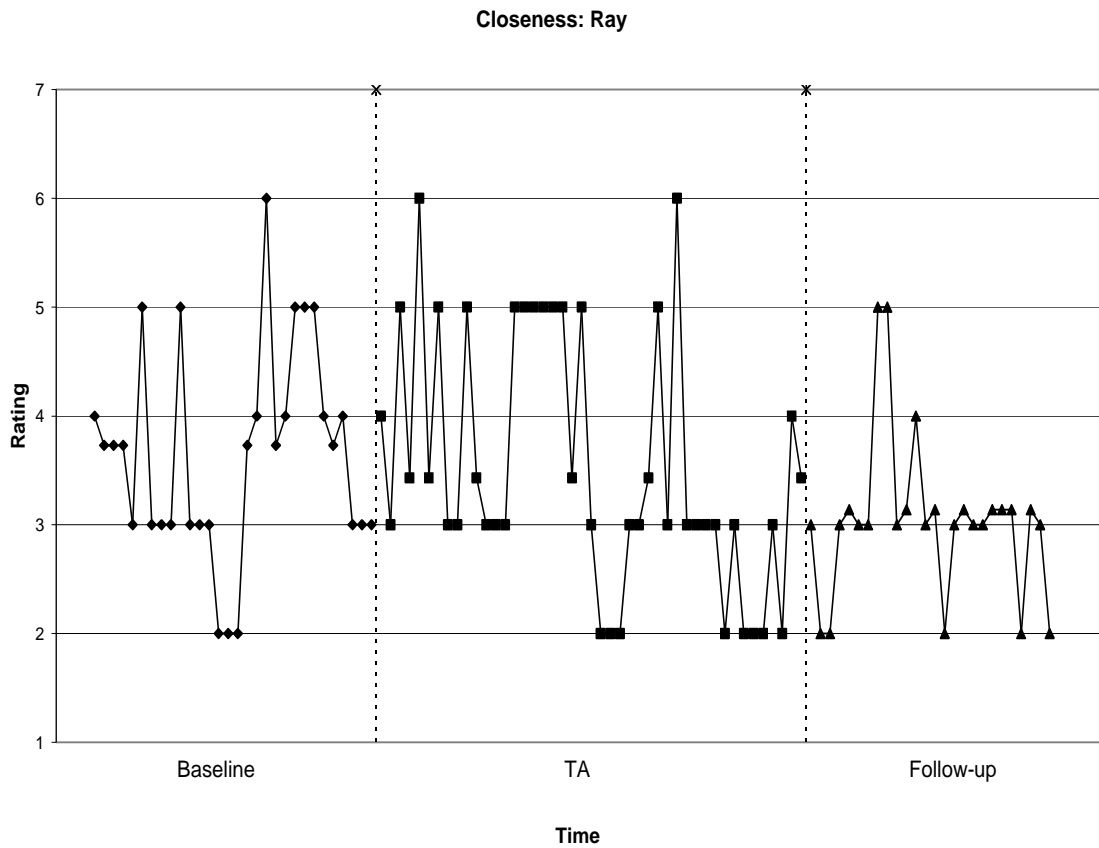
The means, standard deviations, OLS slopes, and number of data points for each of Ray’s daily questions are presented in Table A1 (Appendix J). For each daily question, phase effect analyses were run to determine whether there was a significant change in the slope and level of the scores from the baseline to the TA phase, and from the baseline to the follow-up (FU) phase. The estimation maximization procedure (EM) was used to estimate missing data values. Data were missing for 17.1% of Ray’s daily questions. The MCAR statistic was unavailable due to insufficient degrees of freedom for the chi squared test.

Results of the phase effect analyses, along with line graphs of Ray’s ratings on each of the daily questions are presented below. Significant ($p < .05$) and nearly significant ($p < .10$) results are listed in **bold**.

2. How close did you feel to Mary today? (1=Extremely close, 7=Extremely distant)

- Baseline vs. TA phase: baseline mean (3.65) > TA mean (3.55). $r(\text{Lag } 1) = .34$.
 - No significant level change: $r = -.05$; $p = .76$.
 - **Significant slope change: $r = .33$; $p = .04$.**
- Baseline vs. FU phase: baseline mean (3.65) > FU mean (3.04). $r(\text{Lag } 1) = .44$.
 - No significant level change: $r = -.33$; $p = .11$.
 - No significant slope change: $r = .07$; $p = .76$.

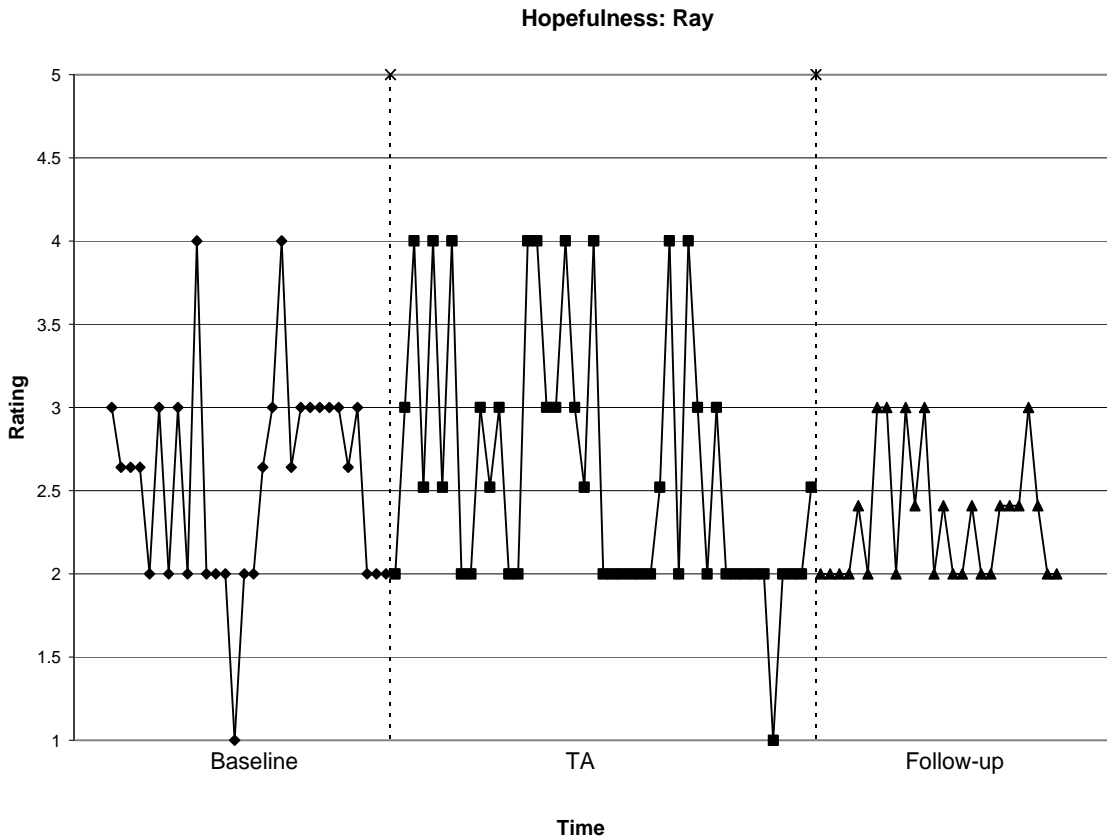
Figure 4.2.



3. How hopeful did you feel about your relationship today? (1=very hopeful, 5=not at all hopeful)

- Baseline vs. TA phase: baseline mean (2.56) < TA mean (2.62). $r(\text{Lag } 1) = .19$.
 - No significant level change: $r = .04$; $p = .78$.
 - **Significant slope change: $r = .27$; $p = .05$.**
- Baseline vs. FU phase: baseline mean (2.56) > FU mean (2.32). $r(\text{Lag } 1) = .20$.
 - No significant level change: $r = -.02$; $p = .18$.
 - No significant slope change: $r = -.22$; $p = .90$.

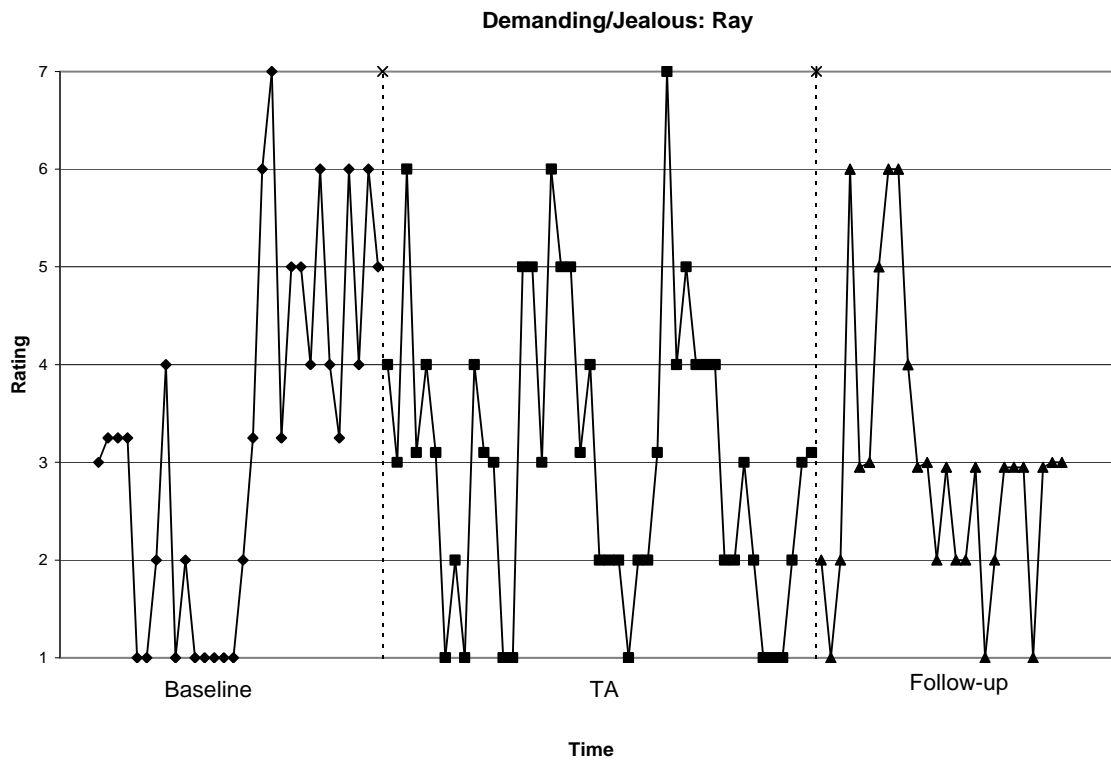
Figure 4.3.



4. On a scale of 1-7, how much were you bothered today by the feeling that Mary was demanding too much of your time or being jealous of the time you spent with others? (1=not at all bothered, 7=extremely bothered)

- Baseline vs. TA phase: baseline mean (3.28) > TA mean (3.06). $r(\text{Lag } 1) = .54$.
 - No significant level change: $r = -.07$; $p = .75$.
 - No significant slope change $r=.32$; $p=.12$.
- Baseline vs. FU phase: baseline mean (3.28) > FU mean (2.98). $r(\text{Lag } 1) = .55$.
 - No significant level change: $r = -.09$; $p=.73$
 - **Significant slope change: $r=.45$; $p=.05$.**

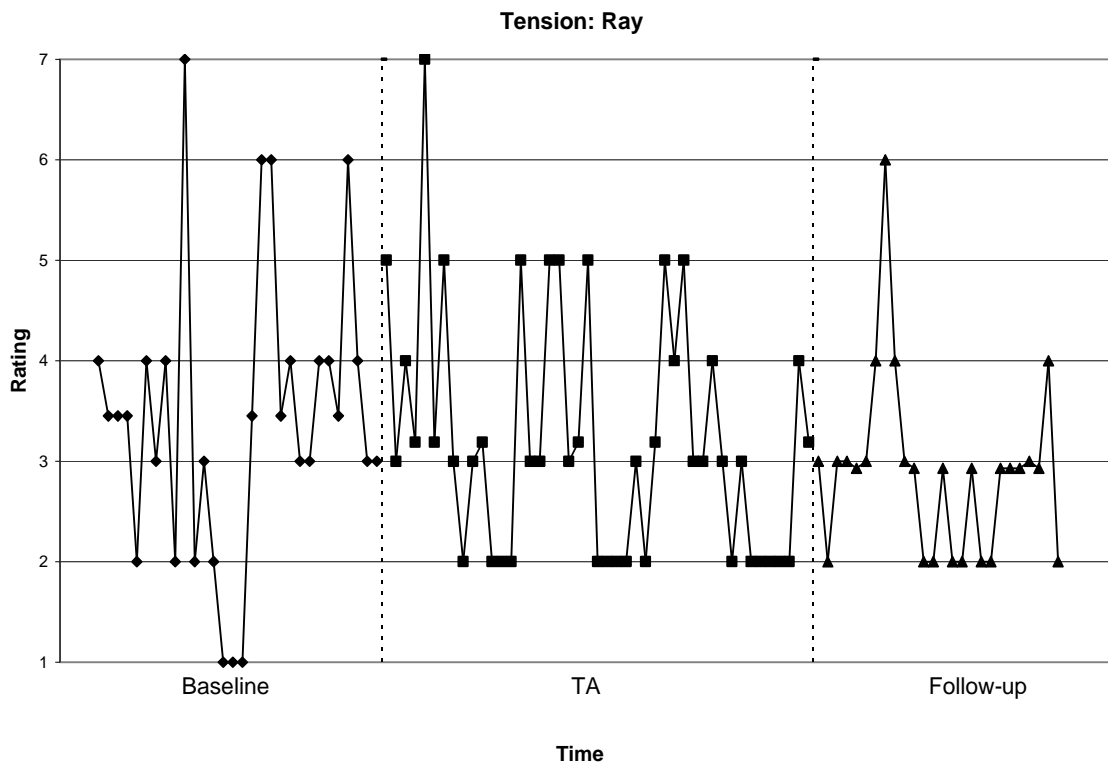
Figure 4.4.



6. Please rate the degree of tension between you and Mary today. (1=no tension, 7=extreme tension)

- Baseline vs. TA phase: baseline mean (3.42) > TA mean (3.23). $r(\text{Lag } 1) = .20$.
 - No significant level change: $r = -.07$; $p = .62$.
 - **Nearly significant slope change: $r = .24$; $p = .09$.**
- Baseline vs. FU phase: baseline mean (3.42) > FU mean (2.90). $r(\text{Lag } 1) = .29$.
 - No significant level change: $r = -.21$; $p = .24$.
 - No significant slope change: $r = .12$; $p = .51$.

Figure 4.6.

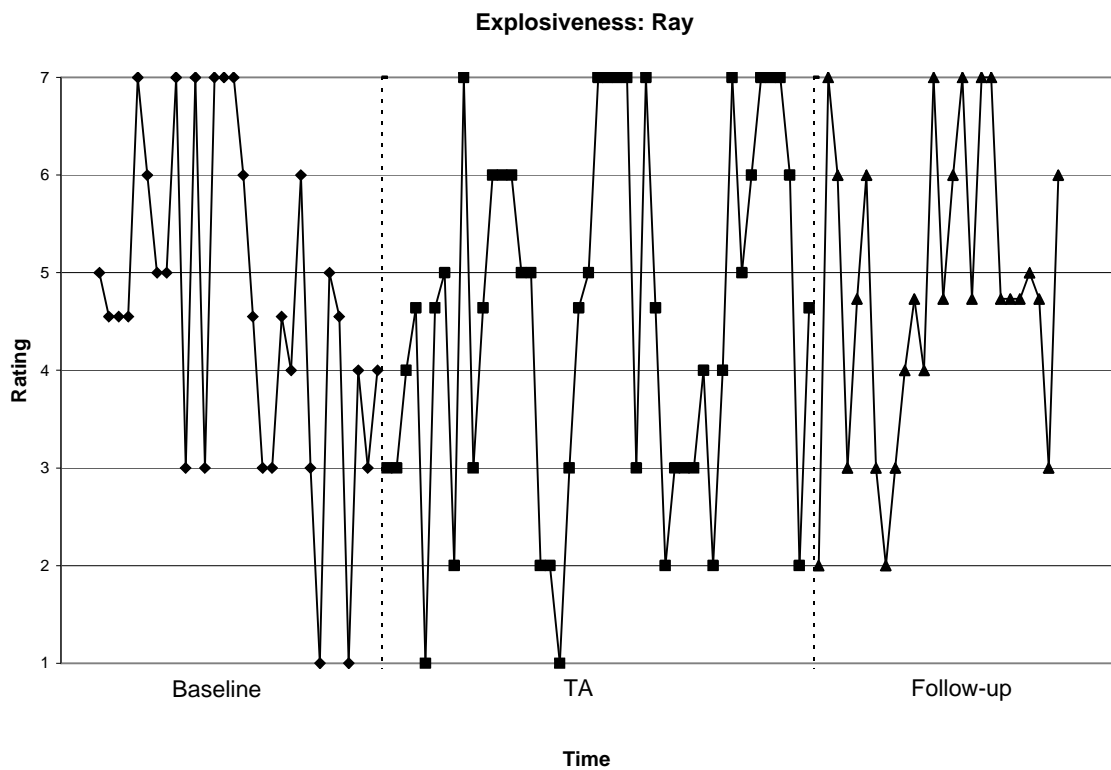


7. Please rate the degree of explosiveness in your relationship with Mary today.

(1=extremely explosive, 7=not at all explosive).

- Baseline vs. TA phase: baseline mean (4.64) > TA mean (4.51). $r(\text{Lag } 1) = .29$.
 - No significant level change: $r = -.04$; $p = .79$.
 - **Nearly significant slope change: $r = -.26$; $p = .09$.**
- Baseline vs. FU phase: baseline mean (4.64) < FU mean (4.84). $r(\text{Lag } 1) = .16$.
 - No significant level change: $r = .06$; $p = .70$.
 - **Significant slope change: $r = -.35$; $p = .03$.**

Figure 4.7.



Summary of time-series results: Ray

Ray's ratings of his general feelings toward Mary changed in slope, becoming more positive over time when comparing the baseline phase to the TA phase. In addition, the overall level of his general feelings toward Mary improved significantly when comparing the baseline phase to the follow-up phase. He also reported he was significantly more open/transparent in his interactions with Mary when comparing his ratings at baseline to those at follow-up.

An examination of Ray's ratings of the degree to which he felt bothered by feelings that Mary was being "demanding" and "jealous" revealed no significant change in level during the TA or follow-up phases when compared to the baseline phase; however, the significant change in slope during the follow-up phase suggests a trend toward becoming less bothered by these behaviors. There was also a significant change in the slope of his ratings of "explosiveness" between himself and Mary when comparing the baseline phase to both the TA and follow-up phases, indicating a trend toward less explosiveness. Although the level of his overall ratings of interpersonal tension, closeness toward Mary, and hopefulness about the relationship did not improve significantly over time, the change in slope during the TA phase suggests they were beginning to move in a positive direction (trending toward less tension, greater closeness, and increased hopefulness) during this phase; however, the slopes at follow-up were no longer significantly different from the baseline slopes on these three variables.

Time-Series Phase Effect Analyses—Mary

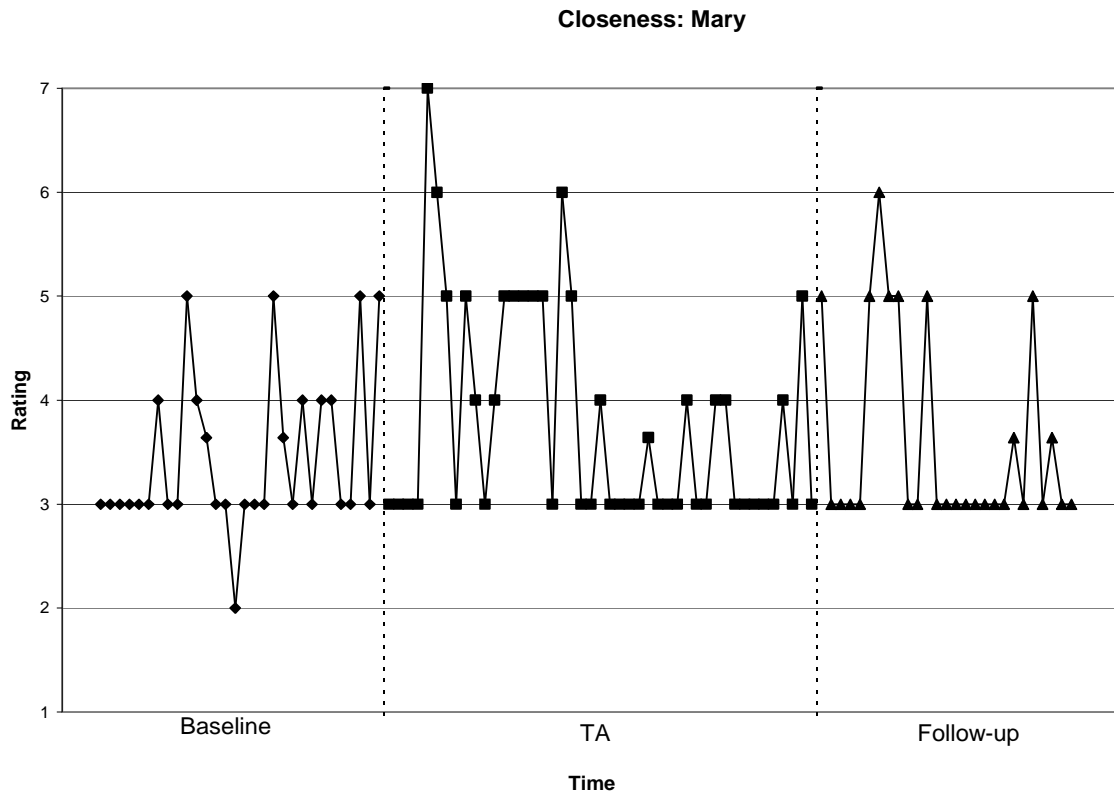
The means, standard deviations, OLS slopes, and number of data points for each of Mary's daily questions are presented in Table A2 (Appendix J). For each daily question, phase effect analyses were run to determine whether there was a significant change in the slope and level of the scores from the baseline to the TA phase, and from the baseline to the follow-up (FU) phase. The estimation maximization procedure (EM) was used to estimate missing data values. Data were missing for 4.9% of Mary's daily questions. The MCAR statistic was unavailable due to insufficient degrees of freedom for the chi squared test.

Results of the phase effect analyses, along with line graphs of Mary's ratings on each of the daily questions are presented in the section below. Significant ($p < .05$) and nearly significant ($p < .10$) results are listed in **bold**.

2. How close did you feel to Ray today? (1=Extremely close, 7=Extremely distant)

- Baseline vs. TA phase: baseline mean (3.44) < TA mean (3.92). $r(\text{Lag } 1) = .26$.
 - No significant level change: $r = .18$; $p = .23$.
 - **Nearly significant slope change: $r = .26$; $p = .09$.**
- Baseline vs. FU phase: baseline mean (3.44) < FU mean (3.60). $r(\text{Lag } 1) = .23$.
 - No significant level change: $r = .09$; $p = .59$.
 - **Nearly significant slope change: $r = .31$; $p = .06$.**

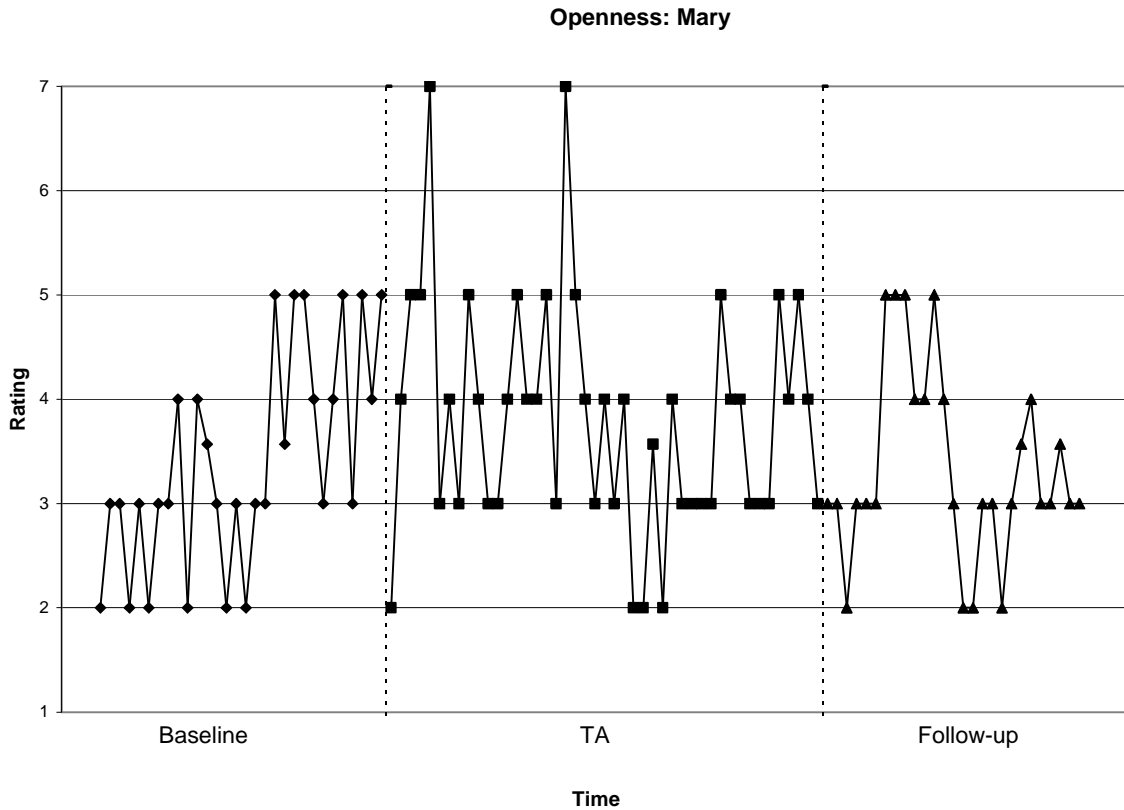
Figure 4.9.



4. On a scale of 1-7, to what degree did you feel Ray was open/transparent vs. private/secretive with you today? (1=very open, 7=very private)

- Baseline vs. TA phase: baseline mean (3.40) < TA mean (3.81). $r(\text{Lag } 1) = .19$.
 - No significant level change: $r = .18$; $p = .20$.
 - **Nearly significant slope change: $r = .23$; $p = .10$.**
- Baseline vs. FU phase: baseline mean (3.40) > FU mean (3.34). $r(\text{Lag } 1) = .41$.
 - No significant level change: $r = .03$; $p = .89$.
 - **Significant slope change: $r = .43$; $p = .02$.**

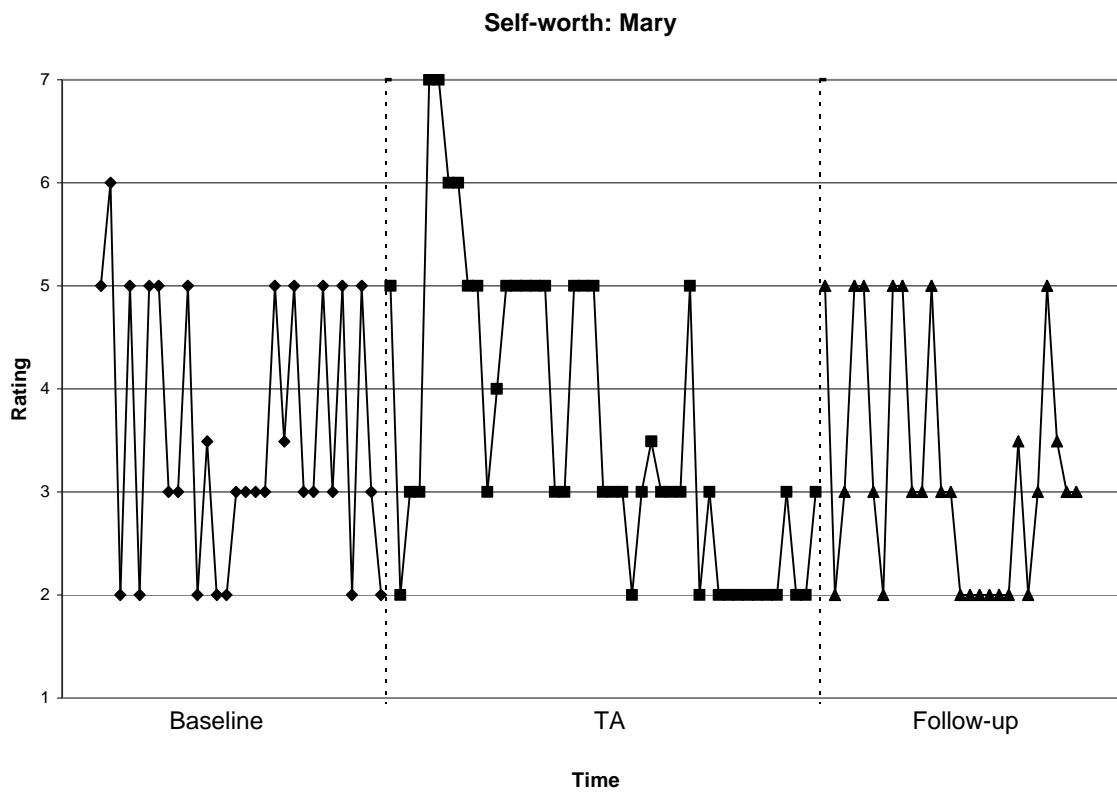
Figure 4.11.



6. How would you rate your own feelings of self-worth or self-respect today? (1=a great deal of self-worth, 7=no self-worth)

- Baseline vs. TA phase: baseline mean (3.57) < TA mean (3.61). $r(\text{Lag } 1) = .33$.
 - No significant level change $r = .02$; $p = .91$.
 - **Significant slope change: $r = .40$; $p = .01$.**
- Baseline vs. FU phase: baseline mean (3.57) > FU mean (3.22). $r(\text{Lag } 1) = .06$.
 - No significant level change: $r = -.14$; $p = .28$.
 - No significant slope change: $r = .02$; $p = .87$

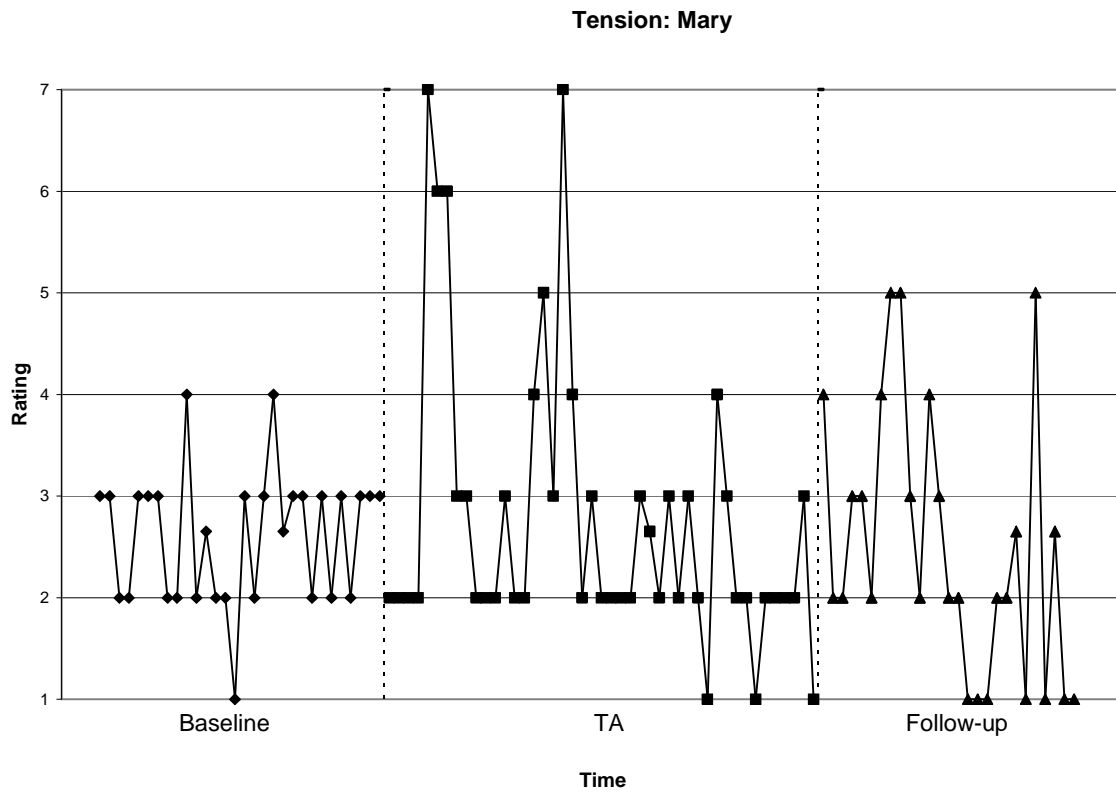
Figure 4.13.



7. Please rate the degree of tension between you and Ray today. (1=no tension, 7=extreme tension).

- Baseline vs. TA phase: baseline mean (2.61) < TA mean (2.77). $r(\text{Lag } 1) = .35$.
 - No significant level change: $r = .07$; $p = .69$.
 - No significant slope change: $r = .27$; $p = .11$.
- Baseline vs. FU phase: baseline mean (2.61) > FU mean (2.49) $r(\text{Lag } 1) = .16$.
 - No significant level change: $r = -.06$; $p = .70$.
 - **Nearly significant slope change: $r = .27$; $p = .09$.**

Figure 4.14.

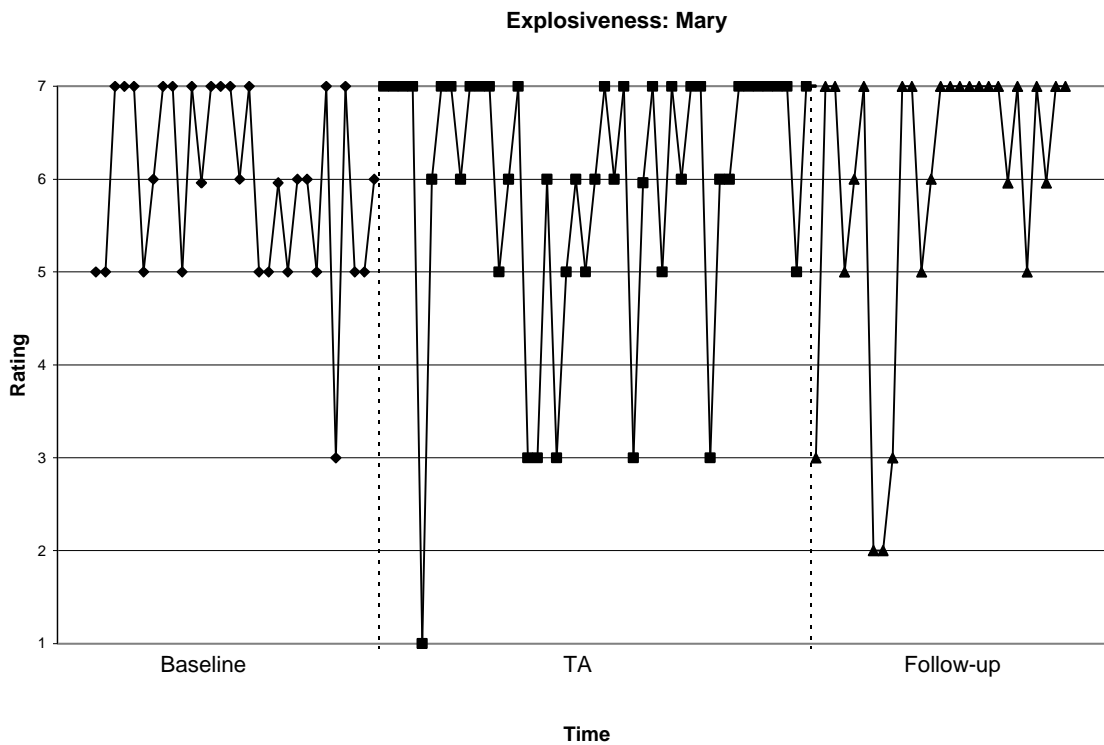


8. Please rate the degree of explosiveness in your relationship with Ray today.

(1=extremely explosive, 7=not at all explosive).

- Baseline vs. TA phase: baseline mean (5.96) = TA mean (5.95). $r(\text{Lag } 1) = 0.0$.
 - No significant level change: $r = -.003$; $p = 1.0$.
 - No significant slope change: $r = -.13$; $p = .27$.
- Baseline vs. FU phase: baseline mean (5.96) = FU mean (5.96). $r(\text{Lag } 1) = .18$.
 - No significant level change: $r = -.002$; $p = 1.0$.
 - **Significant slope change: $r = -.33$; $p = .03$.**

Figure 4.15.



Summary of time series results: Mary

There were no significant changes in the level of any of Mary's daily ratings. However, the slopes of her ratings shifted significantly on several variables, suggesting her ratings on these variables (many of which appeared to be worsening across the course of the baseline phase) were beginning to improve during the TA and/or follow-up phases. There was a significant change in slope on her ratings of hopefulness when comparing the baseline phase to both the TA and follow-up phases, indicating her sense of hopefulness was beginning to increase. The slope of her ratings of Ray's openness vs. secretiveness shifted from the baseline to the TA phase, and this change in slope became statistically significant during the follow-up phase, suggesting she was beginning to view Ray as becoming more open in his interactions with her during and after the TA. There was also a nearly significant change in slope on her ratings of closeness between herself and Ray when comparing the baseline phase to both the TA and follow-up phases, suggesting her feelings of closeness were beginning to improve. There was a significant change in the slope of her ratings of "explosiveness," as well as her ratings of feeling valued/loved by Ray when comparing the baseline to the follow-up phase, indicating she was beginning to experience less "explosiveness" and increased feelings of being valued/loved. There was a nearly significant change in the slope of her ratings of tension with Ray when comparing the baseline to the follow-up phase, suggesting a slight trend toward decreased tension. Regarding her ratings of her own feelings of self-worth, the change in slope during the TA phase suggests they were beginning to move in a positive direction (trending toward greater self-worth) during the phase. These changes in slope, however, were no longer statistically significant at follow-up.

Clients' perceptions of couples therapy

Visual analysis was used to determine whether the TA led to an improvement in Mary and Ray's experience of couples therapy. The means and standard deviations of their Session Evaluation Questionnaire—Client Form (SEQ-C) scores are presented in Table A7 (Appendix K), and the graphs of their weekly SEQ-C ratings are presented in Figures 4.16 through 4.23. Recall that the SEQ-C was used to assess each partner's feelings about each weekly session with their couples therapist, Jane, throughout all three phases of the study. Although the couple was asked to complete weekly SEQ-C ratings during each phase, their reports on this measure were somewhat sporadic due to a few cancelled sessions with Jane. In addition, Mary neglected to complete one of the SEQ-C's during the TA phase and Ray neglected to complete three of the four SEQ-C's during the follow-up phase. Given these missing data points, the results of Mary and Ray's SEQ-C should be interpreted with some caution.

Of note on the graph of Mary's SEQ-C scores are her relatively high ratings on the "smoothness" and "positivity" subscales of the SEQ-C during the last two therapy sessions in the follow-up period, suggesting that, for at least these two sessions, she felt more comfortable and pleased than she felt at any other therapy sessions during the course of the study. In addition, her higher "depth" scores during and immediately after the TA phase suggest she may have experienced these sessions as somewhat more powerful or valuable than the sessions occurring during the other two phases. Ray's scores, by contrast did not appear to undergo any dramatic changes from phase to phase. Due to the fact that he provided ratings for only one of the four therapy sessions in the follow-up phase, it is difficult to determine whether there were notable differences in his perceptions of couples therapy after the TA.

Figure 4.16.

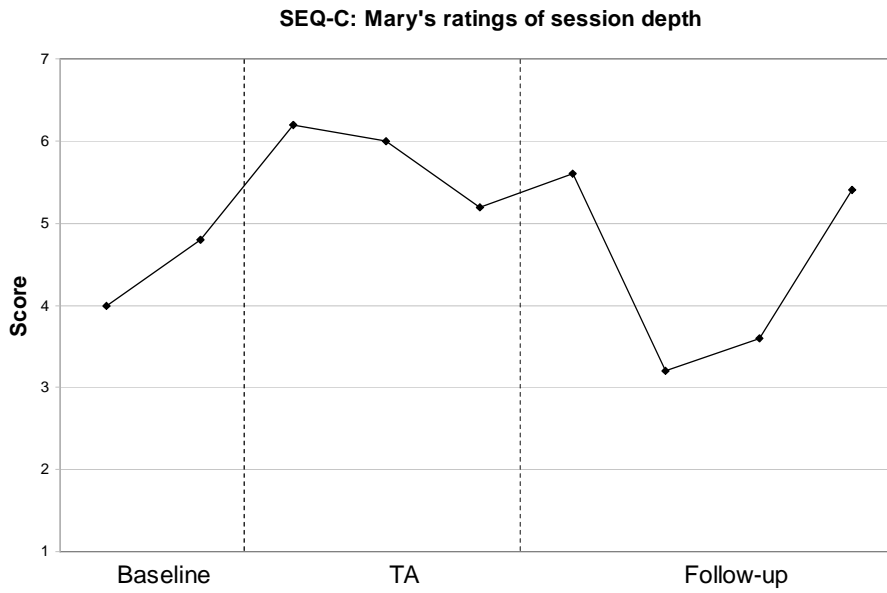


Figure 4.17

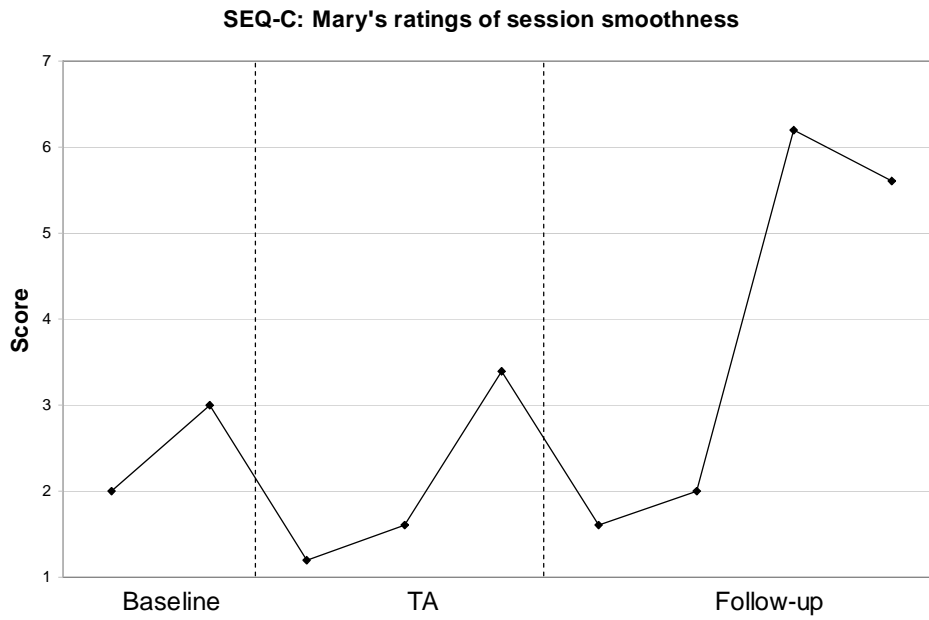


Figure 4.18

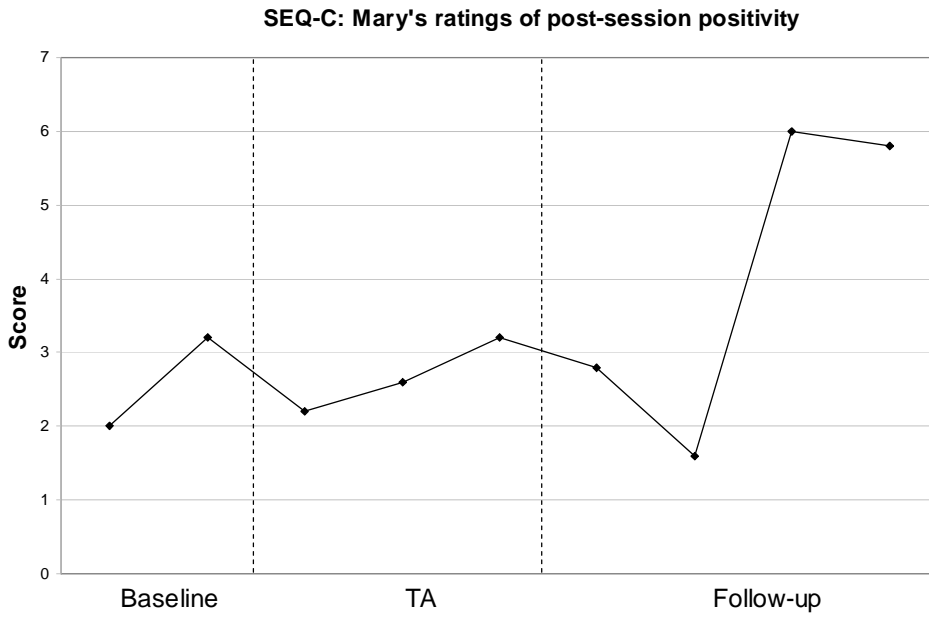


Figure 4.19

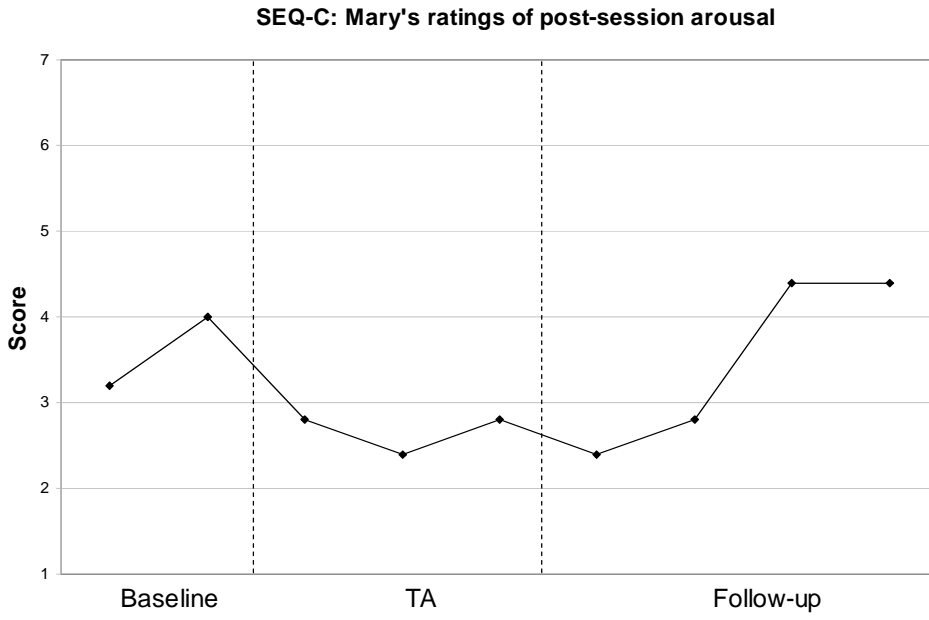


Figure 4.20.

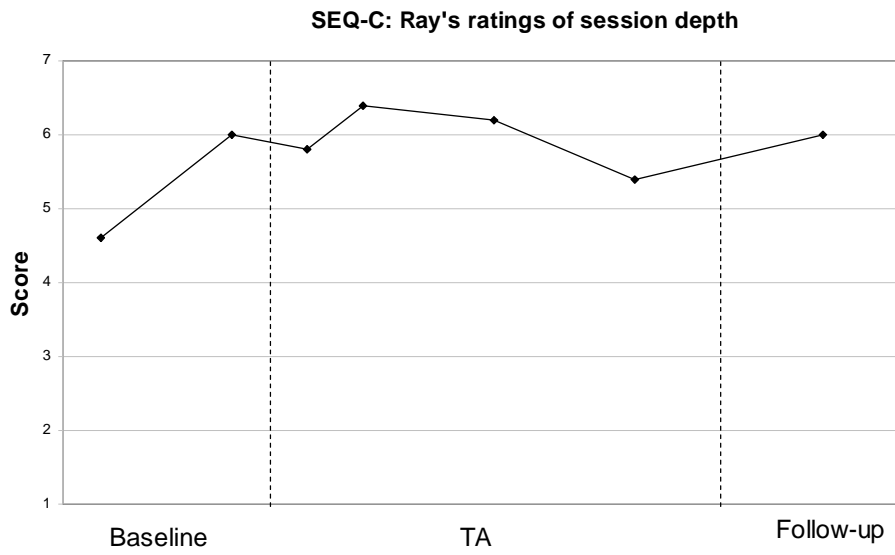


Figure 4.21

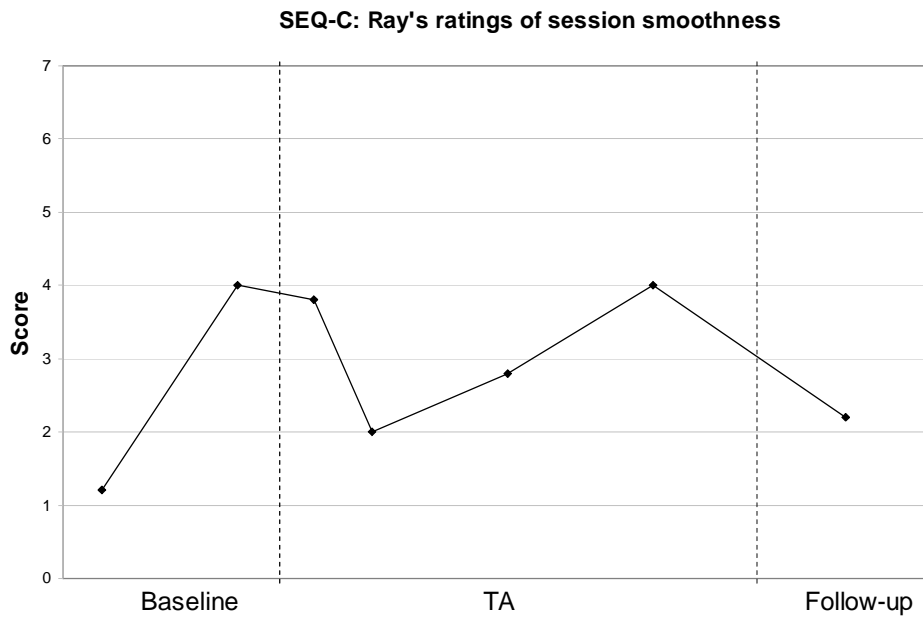


Figure 4.22

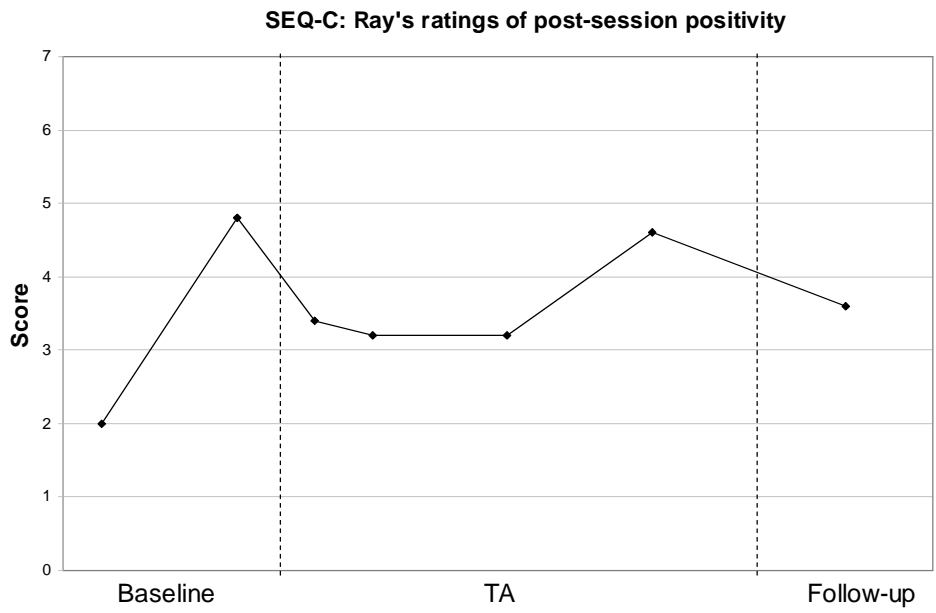
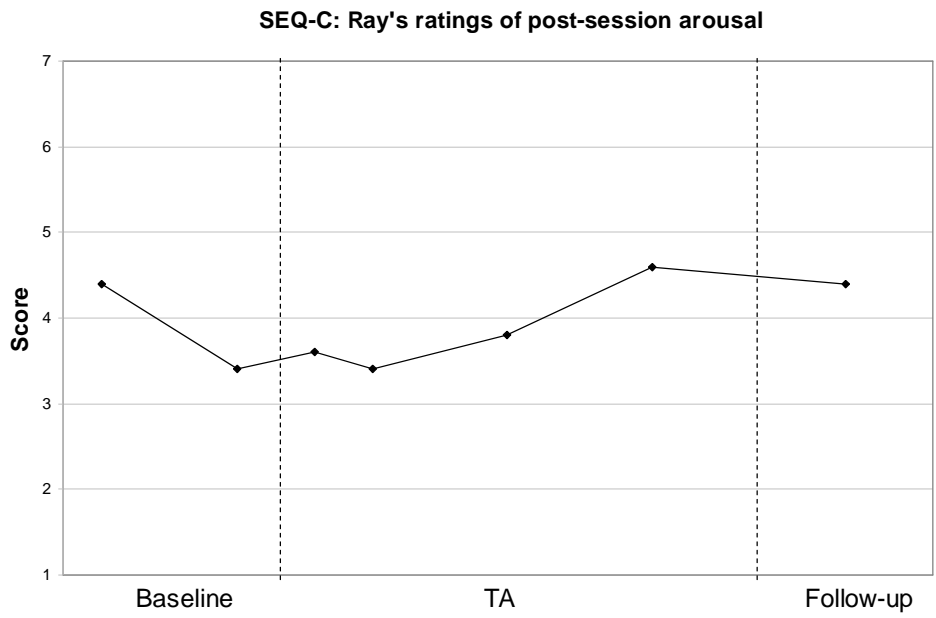


Figure 4.23



Marital satisfaction and psychological symptoms

The couple's BSI, MDS, and DAS ratings are presented in Table A10 (Appendix L), and in Figures 4.24, 4.25, and 4.26. Mary's overall psychological symptoms, as measured by the Brief Symptom Inventory (BSI), decreased by nearly two standard deviations, from the "clinically distressed" range at Time 1 to the "non-distressed" range at Time 2 and Time 3. Ray's BSI, by contrast, increased slightly (by less than half a standard deviation) from Time 1 to Time 3 (he did not complete the Time 2 measures), and both scores fell within the "non-distressed" range.

Mary's marital satisfaction, as measured by the Dyadic Adjustment Scale (DAS), improved somewhat over the course of the study. According to the cutoff scores that are generally used for the DAS (Sabourin, Valois, & Lussier, 2005), her Time 2 and 3 scores fell in the "absent of significant distress range," when compared to her Time 1 score, which fell in the range that indicates neither the presence nor absence of significant marital distress. Ray's DAS scores also improved, from the "distressed" range at Time 1 to the range indicating neither the presence nor absence of significant distress at Time 2.

Mary's scores on the Marital Disaffection Scale (MDS) decreased over the course of the study, indicating she felt lower levels of "disaffection" or "loss of love" for Ray by the end of the study. According to Kayser's (1996) categories, her scores moved from the "below average" to the "low" range. Ray's MDS score, by contrast did not change substantially from baseline to follow-up, although they did move from the "high" to the "above average" range of disaffection.

Figure 4.24.

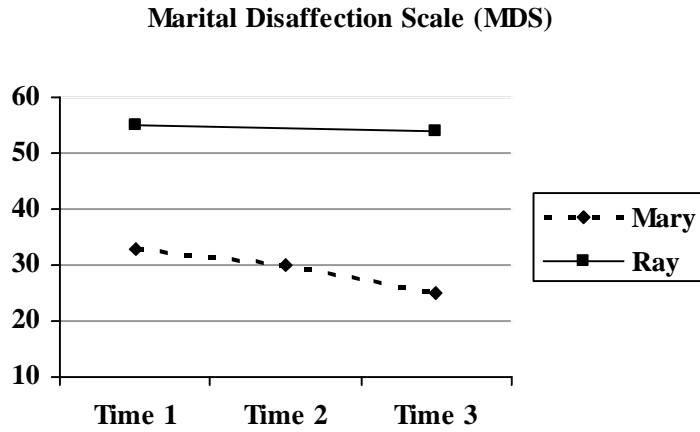


Figure 4.25.

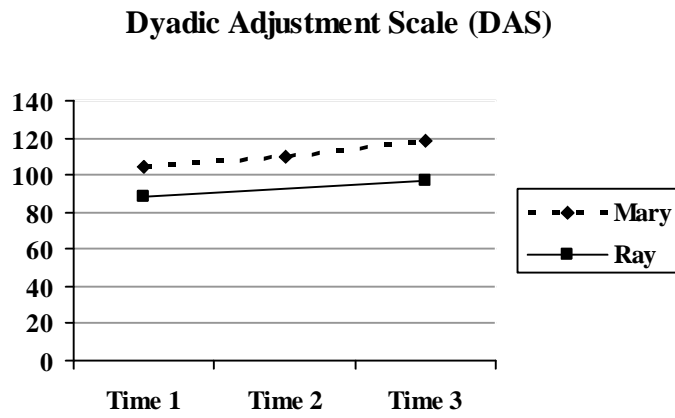
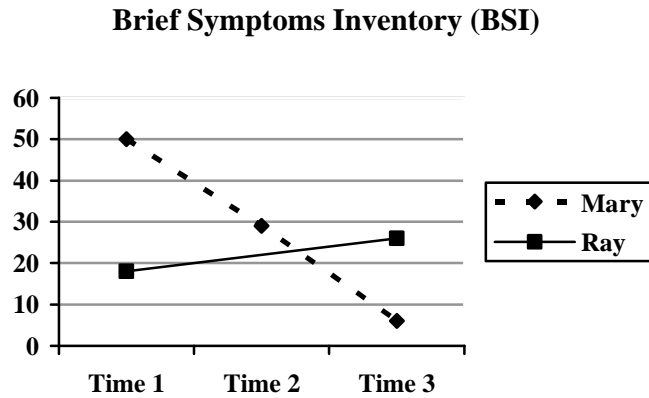


Figure 4.26.



Couple's perceptions of Therapeutic Assessment

Mary and Ray completed the Assessment Questionnaire (AQ-2) once, immediately following the completion of the TA. Their T-scores on each of the subscales of the AQ-2 are presented in Table 4.2. On each of the subscales, both partners' scores fell within half a standard deviation of the mean, indicating their satisfaction with the TA was average when compared to a norming sample of clients receiving psychological assessments at the Center for Therapeutic Assessment.

Table 4.2. *AQ-2 Results (Mean ratings and T-scores)*

	New self-awareness/self-understanding	Positive accurate mirroring	Positive relationship with the examiner	Negative feelings about the assessment	Total satisfaction
Mary	Mean=3.92 T=51.1	Mean=3.33 T=48.9	Mean=4.00 T=51.1	Mean=2.50 T=54.4	Mean=3.65 T=50.4
Ray	Mean=4.08 T=52.5	Mean=3.75 T=51.7	Mean=3.67 T=47.4	Mean=2.00 T=51.7	Mean=3.83 T=52.1

Note: Participants rated each item on a 1-5 Likert-type scale, with high scores indicating higher satisfaction, with the exception of the “negative feelings about the assessment” scale, on which higher ratings indicated more negative feelings.

Couple's qualitative feedback about the TA

Mary

Mary's written responses to the open-ended feedback questionnaire were very positive. She reported the TA intervention was “more than what I expected. It was highly informative and helpful.” She stated the most valuable part of the assessment was identifying “what my strengths and weaknesses are and how I interact with people, particularly, my husband.” She said she could not think of any “least valuable” aspects of the assessment. When asked how the assessment affected her relationship with Ray, she replied, “I have a better understanding of my husband and of myself in how I relate to him. I've realized how I act greatly impacts his emotions and feelings, more than I ever thought before.” When asked how the assessment affected her experience in couples therapy, she replied that it “confirms my love for my husband and how devoted I am to

our marriage and family. Therapy helps in making our relationship stronger.” When asked what she would tell a friend considering a TA, she replied, “That it helps in more ways than just your marriage. I gained so much knowledge about who I am and what I need to work on in me to become the happy person I want to be for my family.” In her final comment, she said the following:

Although, there were times that were incredibly painful during this whole process, I am so thankful that we had this opportunity to take part in this project. The timing couldn't have been better. I learned so much about my husband and myself and feel that with that knowledge we will move forward in our marriage and our love for one another has only become stronger and deeper.

Ray:

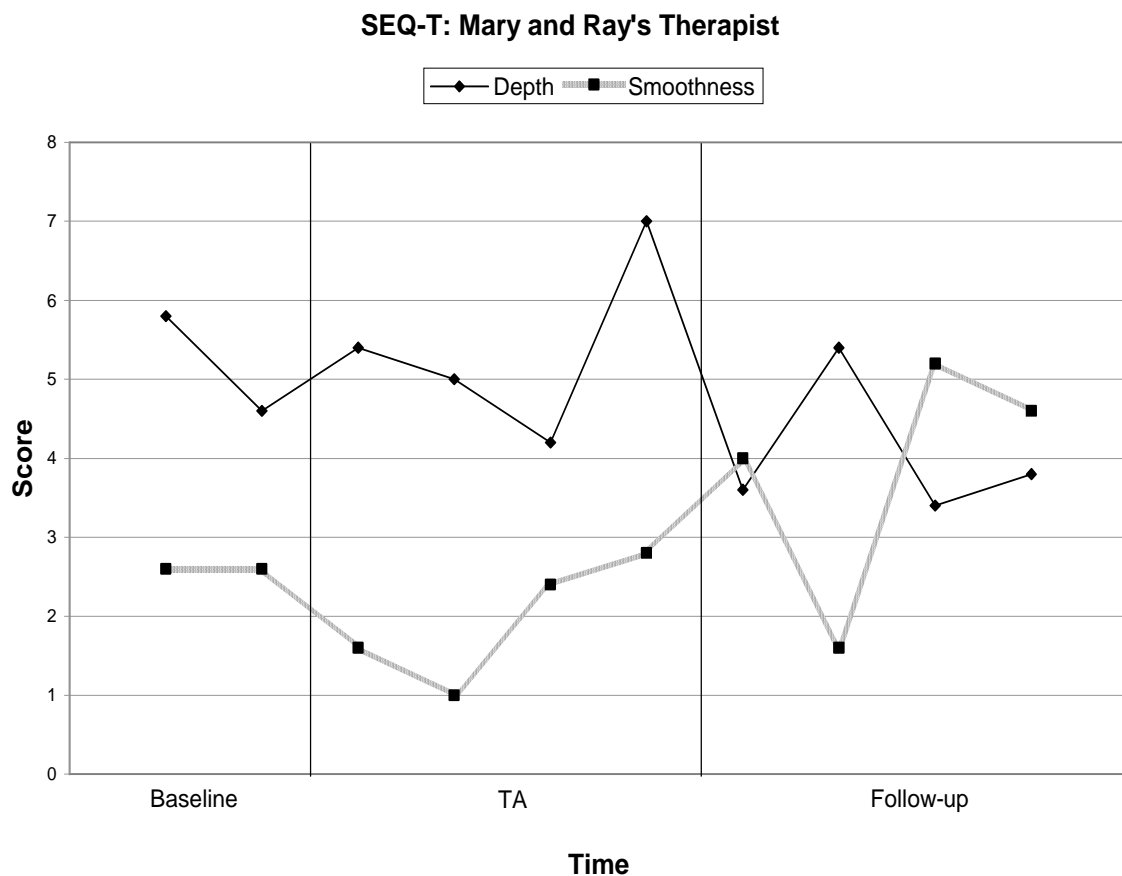
Ray’s written feedback about the TA was largely positive. He reported, “I wasn’t sure what to expect, but I liked the process and outcome” of the Therapeutic Assessment. He reported the most valuable parts of the assessment were the “descriptions about how we interact with others and why.” He stated he could not think of any “least valuable” parts. When asked how the assessment affected his relationship with his partner, he replied, “We are a little more introspective about how are [sic] past shapes our current behavior toward each other.” He also reported the assessment “added additional information and 'material' to assist us in our [couples] therapy.” When asked for suggestions on how the Center for Therapeutic Assessment might improve the way they do assessment, Ray replied, “Maybe more upfront information about the process and goals of the assessment.” When asked what he would tell a friend who was considering getting an assessment, he replied, “That it is a non-intimidating safe way to dig down and discover what is driving your personal behavior and how you interact with partner.”

Therapist’s ratings of couples therapy—Mary and Ray

To ascertain Mary and Ray’s therapist’s perceptions of couples therapy before, during, and after the TA, her weekly ratings on the Session Evaluation Questionnaire—

Therapist Form (SEQ-T) were examined. The means and standard deviations of her SEQ-T scores are presented in Table A13 (Appendix M) and a chart of her weekly SEQ-T ratings is presented in Figure 4.27. For Mary and Ray’s therapist, the “deepest” therapy session occurred near the end of the TA period, while the “smoothest” sessions occurred at the end of the follow-up period. These scores in some way mirror Mary’s ratings, suggesting that both Mary and her therapist were perceiving the sessions similarly.

Figure 4.27.



Therapist’s qualitative feedback—Mary and Ray

Mary and Ray’s therapist’s written feedback about the TA was largely positive. She listed the following as her “hopes and expectations” when she referred the couple for a psychological assessment: “1. Help them increase empathy for each other to decrease

conflict and escalation of conflict. 2. Increase their motivation to engage in individual therapy. 3. To increase their awareness re: how they are vulnerable to reacting and how they set each other off.” She reported that the TA met her expectations in all areas, except in increasing their motivation to engage in individual therapy. When asked about the most useful part of the assessment, she replied that it helped increase “their insight towards each other,” and also “increased compassion and empathy for one another.” Although she did not identify any part of the assessment as being “least useful,” she did note that the couple “may have been somewhat overwhelmed by the intensity of the whole experience.” She reported she could not think of anything that would have made the assessment more useful, and stated, “I would highly recommend this type of assessment, without reservation, to any client or professional.”

Discussion: Mary and Ray

When interpreting Mary and Ray’s results, it is important to consider the disclosure of Ray’s affair that occurred early in the TA phase. Several questions arise around this event and how it relates to the TA intervention: Did their involvement in the TA help the couple get to a point where Ray felt “safe” enough to disclose the affair? Did the TA help the couple recover from this substantial blow to their relationship? Would Ray have disclosed the affair if they had *not* been participating in the TA, and would the manner in which the couple coped with this event have been different? Although it is impossible to know for sure, the improvement on many of the dependent variables suggests the TA may have provided the couple with some degree of help during this crisis. In addition, Mary’s comment that the timing of the TA “couldn’t have been better” suggests the intervention may have provided the couple with resources and/or support that helped them work through this event and its aftermath in an adaptive manner.

Another interesting finding involves the couple’s BSI scores, which revealed Mary’s reported level of symptomatology decreased substantially from baseline to follow-up, while Ray’s symptoms increased slightly (but still remained in the “non-distressed” range). One possible explanation for this finding is that the TA may have had

a sort of balancing effect on the couple. Middelberg (2001) noted that couples will often perform a “dance” in which the partners

assume the polarized positions of the overresponsible caretaker/parent and the underresponsible patient/child...When the symptomatic partner starts to get better and starts to be more functional and self-responsible, the overresponsible partner’s self-image as the protector/rescuer is threatened in that the needs for dependency can no longer be projected onto the partner. This creates anxiety (p. 347).

Although one can only speculate based on the couple’s BSI results, it is possible that the change in scores reflects the kind of “dance” Middelberg describes.

On the whole, Mary and Ray’s results look promising. They showed small improvements on their standardized measures of marital satisfaction, and their time-series results showed that they were at least moving in the right direction on many of their presenting problems. In addition, Mary and Jane’s SEQ ratings suggest the TA may have led to somewhat “deeper” couples therapy sessions during the TA phase of the study, and later, “smoother” sessions during the follow-up phase. Unfortunately, given that follow-up data were only available for the four-week period after the TA, it is impossible to know at this point whether the couple continued to make improvements in their relationship, and whether the improvements they did make were lasting. Some of this will depend on the degree to which they followed through on recommendations made by the assessor and by their therapist. For example, we do not know whether either partner eventually entered into individual therapy, as Jane had hoped they would. In addition, although the couple made statistically significant improvements on some of their daily measures, it is unclear whether these improvements represent *clinically* significant changes in their level of marital satisfaction. These and other limitations will be discussed in the following chapter.

Couple #2: Susan and George

Susan was in her mid 50s and George in his late 60s at the time of the study. The couple was unmarried, but had been living together for approximately 15 years and had been seeing their therapist, Linda, in couples therapy for three years. George, a white man, had a graduate degree and was retired. Susan, a white woman, had a Bachelor's degree and worked in retail.

The couple had originally presented for treatment because both partners constantly felt frustrated and angry and their arguments had become frequent and explosive. They had also stopped having sex. At the time of the current study, they noted feeling "hugely impressed" with couples therapy; George reported, "It's made a major difference in my life," and Susan said it was "the best personal therapy I've ever had." They noted that couples therapy with Linda had helped them understand ways in which their differing worldviews often contributed to conflict between them. Although therapy had helped them to "recover better and faster" from arguments, they noted that they were still fighting, and that the fighting was "tiring [them] out." Susan noted that they often got into "power struggles," and that a simple question from one of them would leave the other one feeling criticized. In addition, they had not yet seen any improvement in their sex life. Both partners said they would like to find a way to make sex enjoyable for both of them. In addition, Susan said that although she knew George loved her, it was difficult for her to trust his love. She also noted, "I have problems believing my needs will be made equal to his needs" in the relationship. She reported feeling as though she was not entitled to have her emotional needs met by George. She said she also felt as though he was constantly correcting her and challenging her to defend her opinions. All of this often left her feeling angry with him. George, by contrast, felt as though Susan often "talked in code" and would never say what she really wanted. Her lack of directness and transparency often left him feeling frustrated.

After the Time 1 interviews with George and Susan, the researcher used information from the interview, as well as the couple's responses to the DAS, MDS, and BSI, to develop a set of questions to which each partner would respond on a daily basis

throughout the study. The researcher sent drafts of the questions to both partners and invited them to edit or amend the questions as they saw fit. Each of these questions is presented along with the time-series results.

Integrity Rating of TA sessions

Treatment integrity ratings indicated that, in the first TA session, the assessor did not ask Susan and George about their “best guess” answers to their assessment questions, nor did she let them know that she would be asking for their input and collaboration throughout the process. In addition, she did not have the couple re-phrase a non-systemic question that Susan asked about George. In all other areas, however, the assessment was conducted according to the TA principles described by Finn (2007) (see Table 4.3).

Table 4.3. *Treatment integrity ratings—Susan and George*

Initial session with couple	1 Strongly Disagree	2 Disag ree	3 Neither agree nor disagree	4 Agree	5 Strongl y Agree
1. The assessor sought assessment questions from both clients.					X
2. The assessor asked the clients about their “best guess” answers to at least one of these questions.		X			
3. The assessor let the clients know that he/she would be asking for their input/collaboration throughout the assessment process.	X				
4. The assessor talked to the clients about the tests they would be taking.					X
5. The assessor demonstrated empathy.				X	
6. The assessor demonstrated good listening skills.				X	
7. The assessor encouraged the clients to ask questions about the assessment.					X
8. If a client asks a non-systemic question about his/her partner, the assessor helped to rephrase this as a systemic question.		X*			

Table 4.3, cont.

Consensus Rorschach/Assessment Intervention session	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. During the Consensus Rorschach task, the assessor asked the couple to come up with responses that they could both see and agree on.					X
2. After administering the Consensus Rorschach, the assessor asked the clients to talk about their experience during the task (e.g., thoughts, feelings, observations).					X
3. The assessor attempted to make connections between what happened during the Consensus Rorschach task and what happens in other areas of the couple's life together.					X
4. The assessor asked the couple to try out a new approach to the Consensus Rorschach (using a card that was not used during the first administration).					X
5. The assessor demonstrated empathy.					X
6. The assessor demonstrated good listening skills.					X
7. The assessor made systemic interpretations about how each person contributes to the patterns revealed.				X	
Summary/Discussion session	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. The assessor appeared to begin by presenting Level 1 findings, followed by Level 2 and (if appropriate) Level 3 findings.				X	
2. The assessor encouraged the clients to revise or amend the findings.					X
3. The assessor demonstrated empathy.					X
4. The assessor demonstrated good listening skills.					X
5. The assessor seemed to give equal emphasis to each partner's contribution to the relationship struggles.					X
6. The assessor asked clients to give real life examples of assessment findings that were discussed.				X	

*Rater noted, "I have this rating because the woman suggested questions for her partner, phrased from his point of view, about him, and these were not reframed systemically. Although he accepted the questions when she suggested them, it seemed to me that they were really more her questions about him."

Results: Susan and George

Qualitative information from daily surveys

The following prompt was included on each of the couple's daily surveys: "If anything happened today that you think affected your feelings about your relationship in any way, feel free to mention it in the box below." To provide a sense of some of the situational factors that may have affected or been affected by the couple's experience of the TA intervention, their qualitative responses to this prompt will be summarized here.

Throughout the baseline phase, Susan used the comments section of the daily survey to describe both conflicts and moments in which she felt comforted by George or enjoyed his company. For example, in the third week of the baseline phase, she wrote, "He didn't finish the lawn work this week when he said he would. I felt like I couldn't rely on him." On the second-to-last day of the baseline phase, she wrote, "I wasn't feeling well and had to work. He was very attentive and forgiving." Similarly, George described a mixture of positive and negative events related to their relationship throughout this phase. For example, in the fourth week of the baseline phase, he wrote, "A bad spell and power struggle caused as almost always by a communications breakdown. We did see it and talk through it. There is anger not far from the surface which comes up for Susan as anger and for me as depression." Three days later, he wrote, "We worked on something together and treated each other quite well (she treated me well and I think I did ok)."

Both partners reported similar types of ups and downs throughout the TA phase. For example, in the second week of the phase, Susan wrote, "He is still not feeling well. So I feel that burden that I have to be the caretaker adult. I am feeling resentful that he doesn't get things done when he is feeling well and now he is sick so he gets a pass." Approximately one week later, she reported, "I didn't feel well and he took very good care of me." In several of George's comments during this phase, he indicated he believed he and Susan were improving in their ability to resolve conflicts with one another. For example, during the second week of the TA phase, he wrote, "Morning negative event. We dealt w/ it pretty well--we are getting better."

Both partners reported two arguments during the follow-up phase, and otherwise, did not use the comments section of the survey during this phase. The first of these occurred on the first day of the follow-up phase. Then, approximately two weeks into the follow-up phase, both partners described an argument that seemed to stand out as particularly hurtful to both of them. Susan wrote, “This morning I got upset over something that happen[ed] at work and wanted him to comfort me. The more he tried the more upset I got. I could not re settle myself. He got very angry and despairing. He is still angry with me.” On the same day, George wrote, “Another event like Fri 2 wks ago, but worse for me. It's like dealing w/ my mother at one level, and like Susan has to have a fight at another. One or two more and I am GONE.” Three days later, George wrote, “[today's] session w/ [Linda, couples therapist] made me more hopeful.”

Time Series Phase Effect Analyses: George

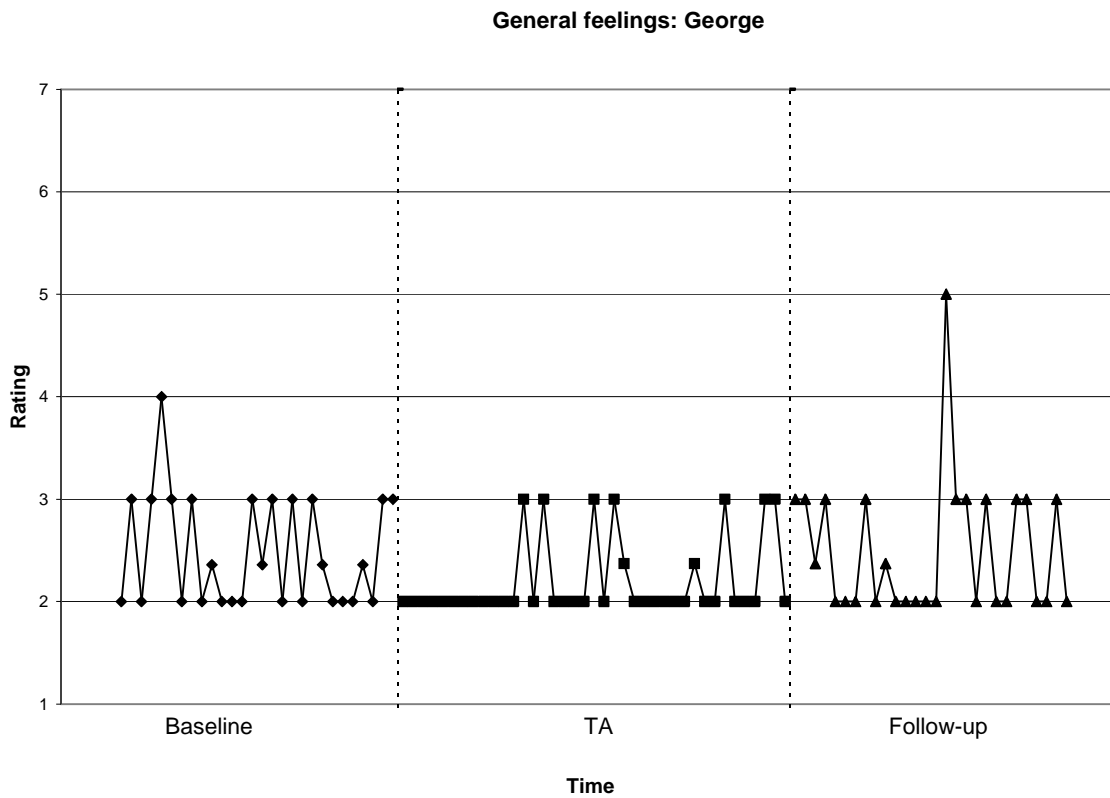
The means, standard deviations, OLS slopes, and number of data points for each of George's daily questions are presented in Table A3 (Appendix J). For each daily question, phase effect analyses were run to determine whether there was a significant change in the slope and level of the scores from the baseline to the TA phase, and from the baseline to the follow-up (FU) phase. The estimation maximization procedure (EM) was used to estimate missing data values. Data were missing 8.4% of George's daily questions. Little's MCAR test yielded a chi-square of .916 and p-value of .338, indicating the MCAR assumption was met.

Results of the phase effect analyses, along with line graphs of George's ratings on each of the daily questions are presented below. Significant ($p < .05$) and nearly significant ($p < .10$) results are listed in **bold**.

1. Please rate your general feelings toward Susan today (1=Extremely positive, 7=Extremely negative).

- Baseline vs. TA phase: baseline mean (2.48) > TA mean (2.20). $r(\text{Lag } 1) = .04$.
 - **Significant level change: $r = -.29$; $p = .02$.**
 - No significant slope change: $r = -.102$; $p = .44$.
- Baseline vs. FU phase: baseline mean (2.48) = FU mean (2.49) $r(\text{Lag } 1) = .02$.
 - No significant level change: $r = .01$; $p = .94$.
 - No significant slope change: $r = -.5$; $p = .72$.

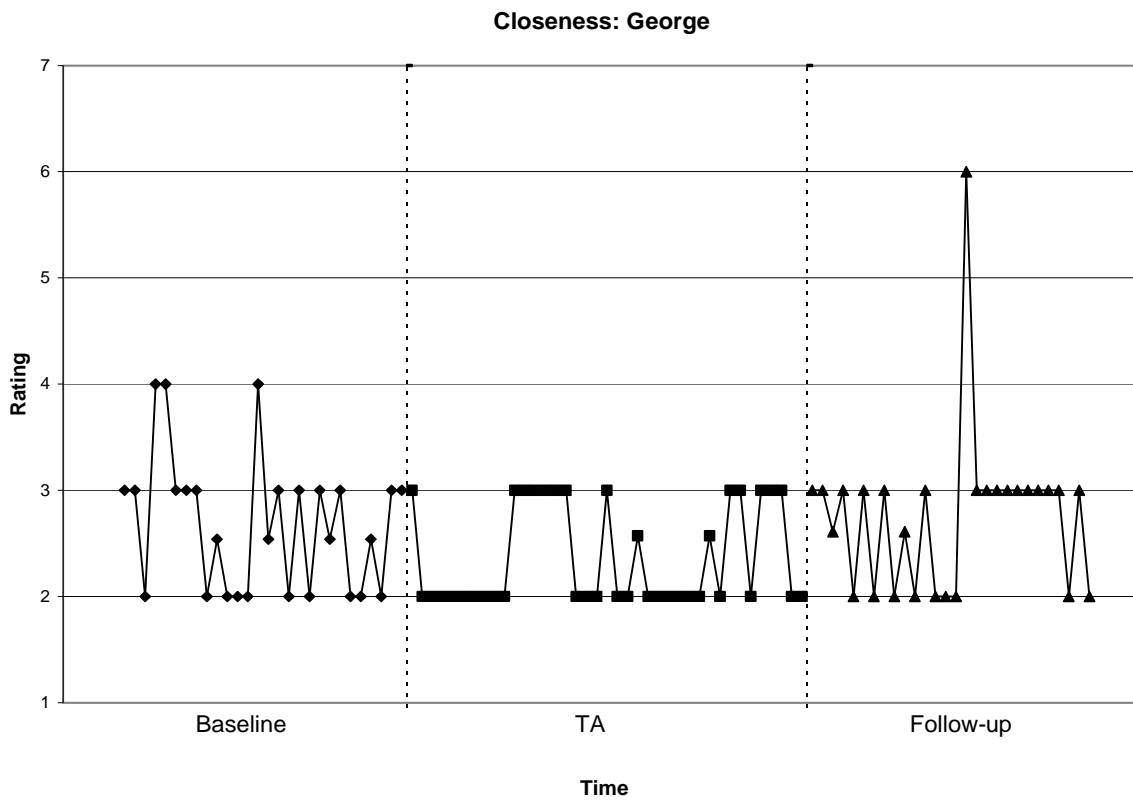
Figure 4.28.



2. How close did you feel to Susan today? (1=Extremely close, 7=Extremely distant)

- Baseline vs. TA phase: baseline (2.68) > TA mean (2.36). $r(\text{Lag } 1) = .26$.
 - **Nearly significant level change: $r = -.28$; $p = .08$.**
 - No significant slope change: $r = -.08$; $p = .62$.
- Baseline vs. FU phase: baseline mean (2.68) < FU (2.76) mean. $r(\text{Lag } 1) = -.04$.
 - No significant level change: $r = .05$; $p = .70$.
 - No significant slope change: $r = -.16$; $p = .22$.

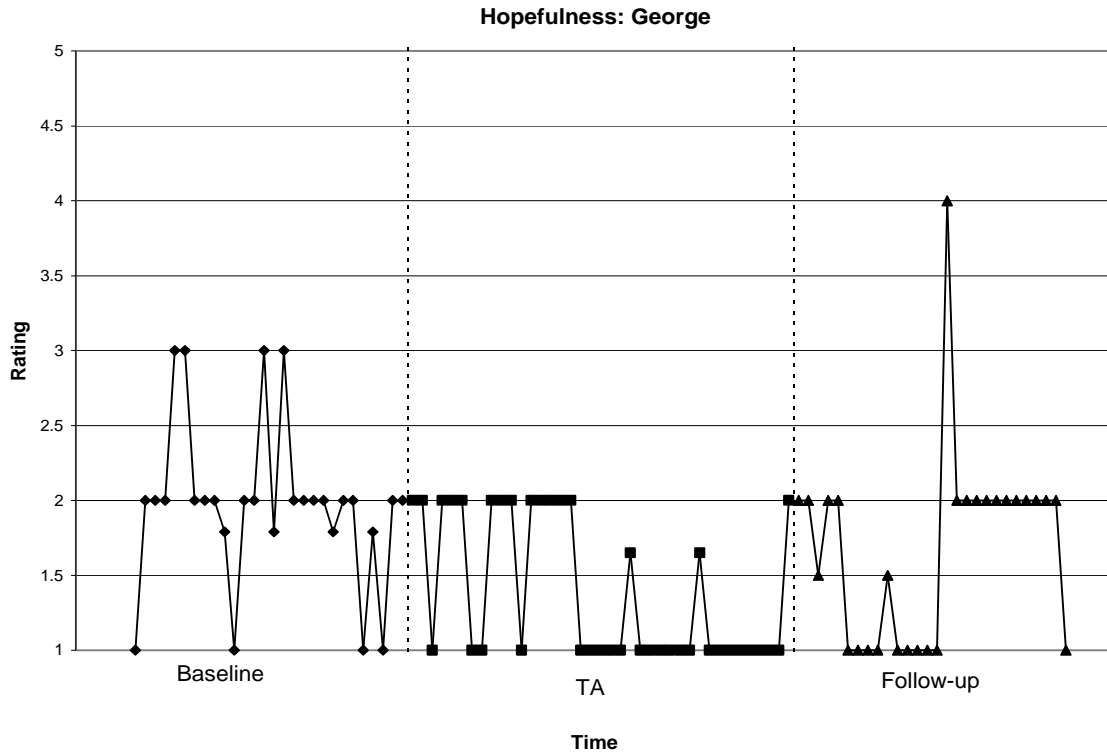
Figure 4.29.



3. How hopeful did you feel about your relationship today? (1=very hopeful, 5=not at all hopeful)

- Baseline vs. TA phase: baseline mean (1.97) > TA mean (1.39). $r(\text{Lag } 1) = .49$.
 - Significant level change: $r = -.49$; $p = .01$.
 - Significant slope change: $r = .34$; $p = .10$.
- Baseline vs. FU phase: baseline mean (1.97) > FU mean (1.68). $r(\text{Lag } 1) = .28$.
 - No significant level change: $r = -.24$; $p = .18$.
 - No significant slope change: $r = -.21$; $p = .24$.

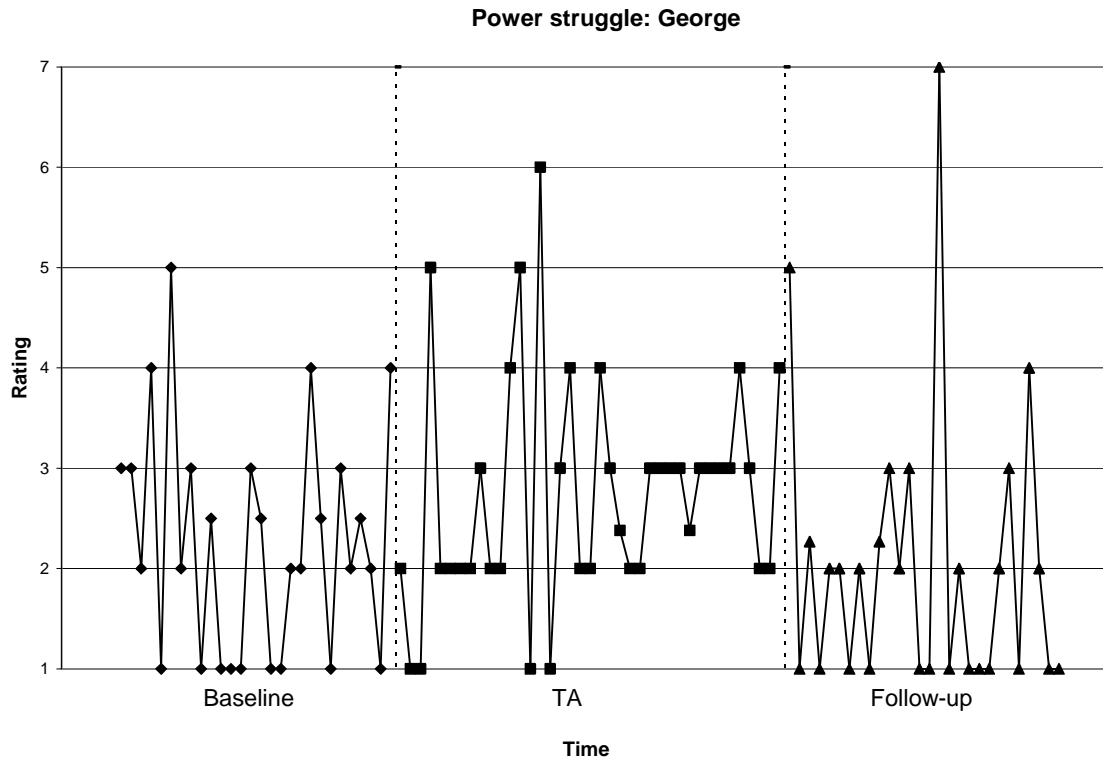
Figure 4.30.



4. On a scale of 1-7, to what degree did you and Susan get caught up in a "power struggle" today? (1 = not at all, 7 = extremely)

- Baseline vs. TA phase: baseline mean (2.25) < TA mean (2.74). $r(\text{Lag } 1) = -.20$.
 - **Significant level change: $r = .21$; $p = .04$.**
 - **Significant slope change: $r = .21$; $p = .04$.**
 - Note that both the level and slope changes indicate *more* power struggles during the TA phase.
- Baseline vs. FU phase: baseline mean (2.25) > FU mean (2.02). $r(\text{Lag } 1) = -.17$.
 - No significant level change: $r = -.09$; $p = .43$.
 - No significant slope change: $r = .01$; $p = .94$.

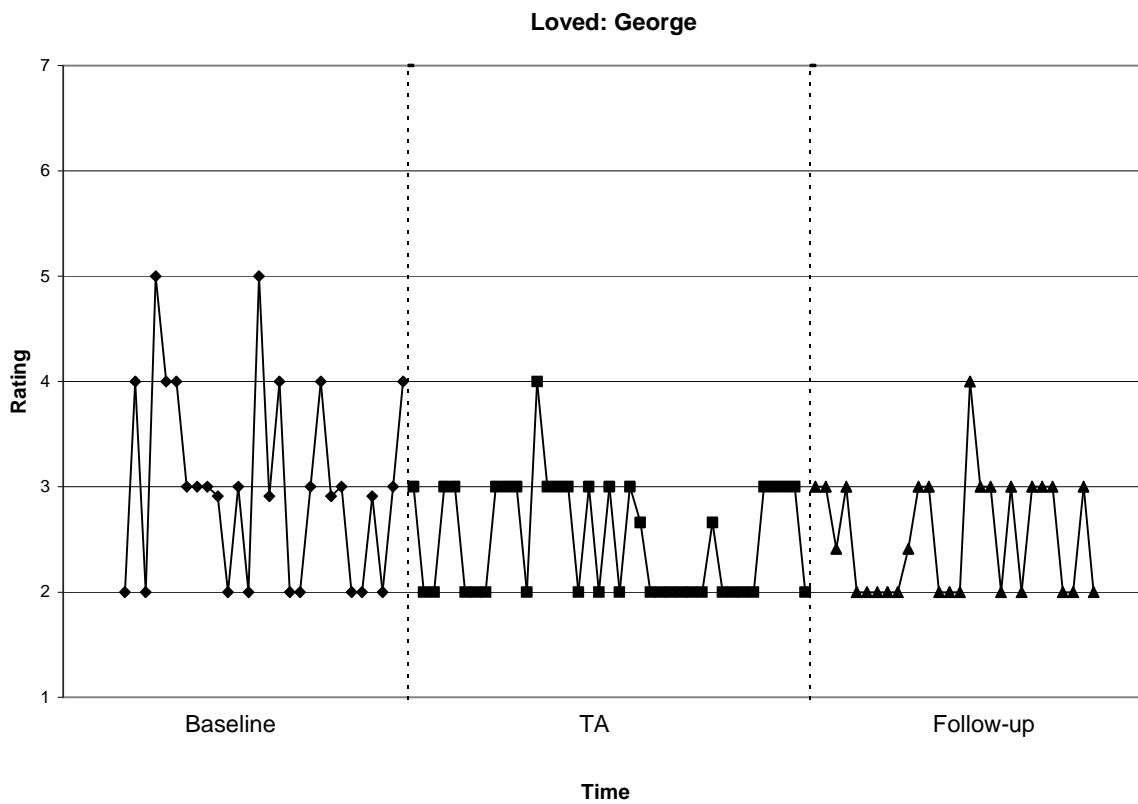
Figure 4.31.



5. On a scale of 1-7, how loved did you feel by Susan today? (1=extremely loved, 7=not at all loved).

- Baseline vs. TA phase: baseline mean (3.02) > TA mean (2.50). $r(\text{Lag } 1) = .11$.
 - **Significant level change: $r = -.34$; $p = .01$.**
 - No significant slope change: $r = .08$; $p = .57$.
- Baseline vs. FU phase: baseline mean (3.02) > FU mean (2.53). $r(\text{Lag } 1) = .06$.
 - **Significant level change: $r = -.31$; $p = .02$.**
 - No significant slope change: $r = -.12$; $p = .41$.

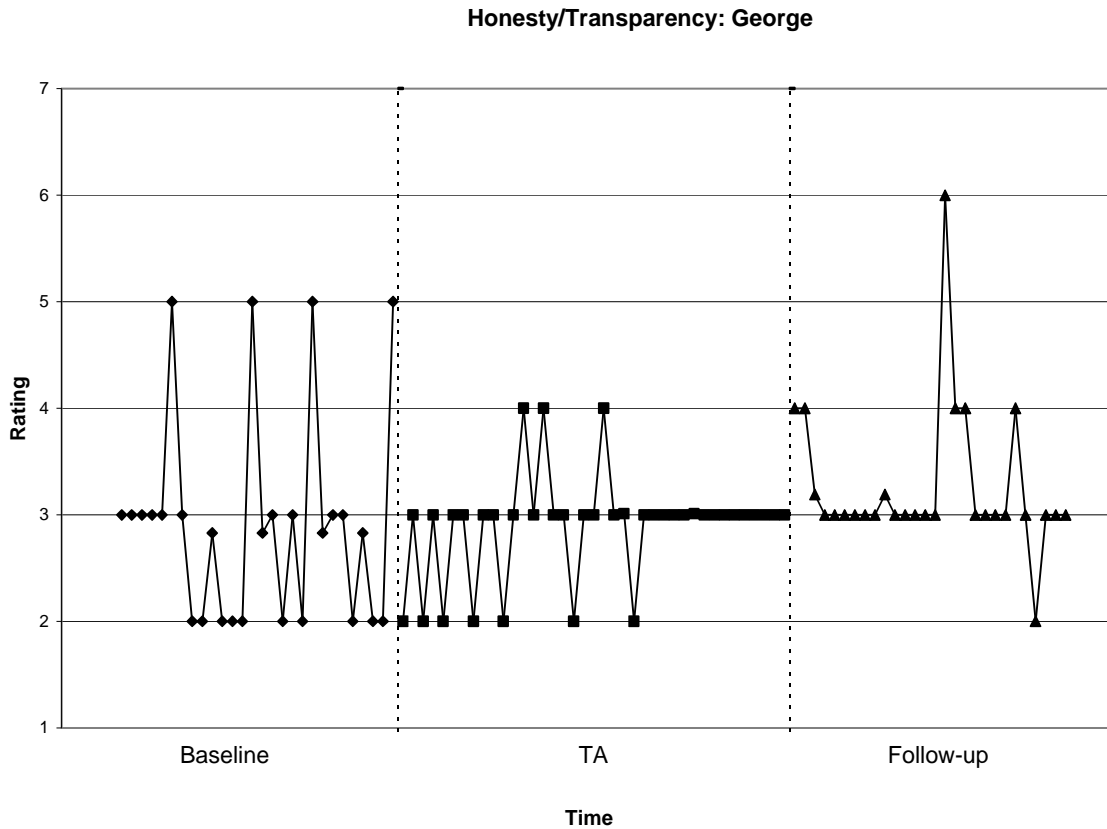
Figure 4.32.



6. In your interactions with Susan today, to what degree did you feel she was being emotionally honest/transparent with you? (1= extremely open; 7=not at all open).

- Baseline vs. TA phase: baseline mean (2.90) = TA mean (2.90). $r(\text{Lag } 1) = -.13$.
 - No significant level change: $r = -.004$; $p = 1.0$.
 - No significant slope change: $r = -.12$; $p = .27$.
- Baseline vs. FU phase: baseline mean (2.90) < FU mean (3.26). $r(\text{Lag } 1) = .12$.
 - No significant level change: $r = .21$; $p = .18$.
 - No significant slope change: $r = .05$; $p = .79$.

Figure 4.33.



7. *Did you and Susan have sex today?*

- No change: his response was “no” on every day of the study.

Summary of time-series results: George

George reported feeling significantly more loved by Susan when comparing the baseline phase to both the TA and follow-up phases. He reported experiencing significantly higher level of hopefulness and more positive feelings toward Susan during the TA phase when compared to the baseline phase; however, his ratings on both of these variables returned to baseline levels during the follow-up phase. Similarly, the improvement in his ratings of closeness toward Susan from the baseline to the TA phase was nearly significant, but again returned to baseline levels during the follow-up phase.

Interestingly, when comparing the baseline to the TA phase, George reported experiencing significantly *more* “power struggles” with Susan, and the significant change in slope suggested these power struggles were increasing somewhat over the course of the phase; however, the slope and level of his ratings on this variable returned to baseline levels during the follow-up phase.

Time Series Phase Effect Analyses: Susan

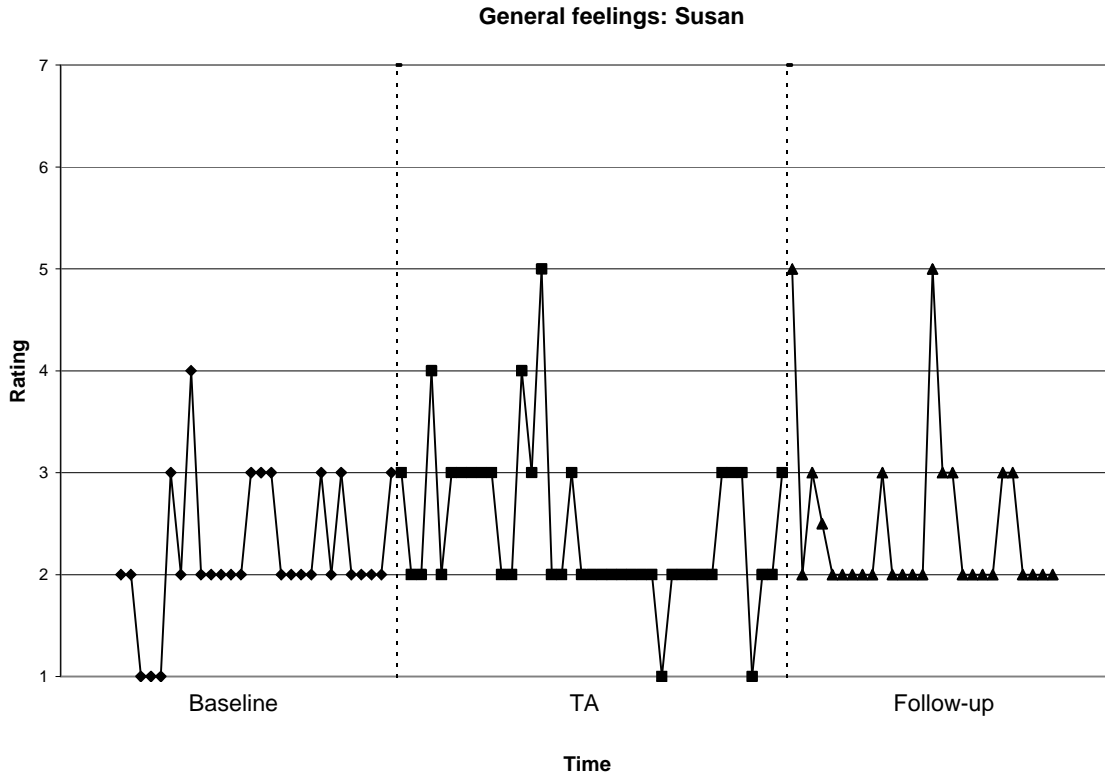
The means, standard deviations, OLS slopes, and number of data points for each of Susan’s daily questions are presented in Table A4 (Appendix J). For each daily question, phase effect analyses were run to determine whether there was a significant change in the slope and level of the scores from the baseline to the TA phase, and from the baseline to the follow-up (FU) phase. The estimation maximization procedure (EM) was used to estimate missing data values. Data were missing for 1.1% of Susan’s daily questions. Little’s MCAR test yielded a chi-square of 1.746 and p-value of .186, indicating the MCAR assumption was met.

Results of the phase effect analyses, along with line graphs of Susan’s ratings on each of the daily questions are presented below. Significant ($p < .05$) and nearly significant ($p < .10$) results are listed in **bold**.

1. Please rate your general feelings toward George today (1=Extremely positive, 7=Extremely negative).

- Baseline vs. TA phase: baseline mean (2.21) < TA mean (2.44). $r(\text{Lag } 1) = .17$.
 - No significant level change: $r = .15$; $p = .31$.
 - **Nearly significant slope change: $r = .25$; $p = .09$.**
- Baseline vs. FU phase: baseline mean (2.21) < FU mean (2.46). $r(\text{Lag } 1) = .16$.
 - No significant level change: $r = .16$; $p = .31$.
 - No significant slope change: $r = .23$; $p = .16$.

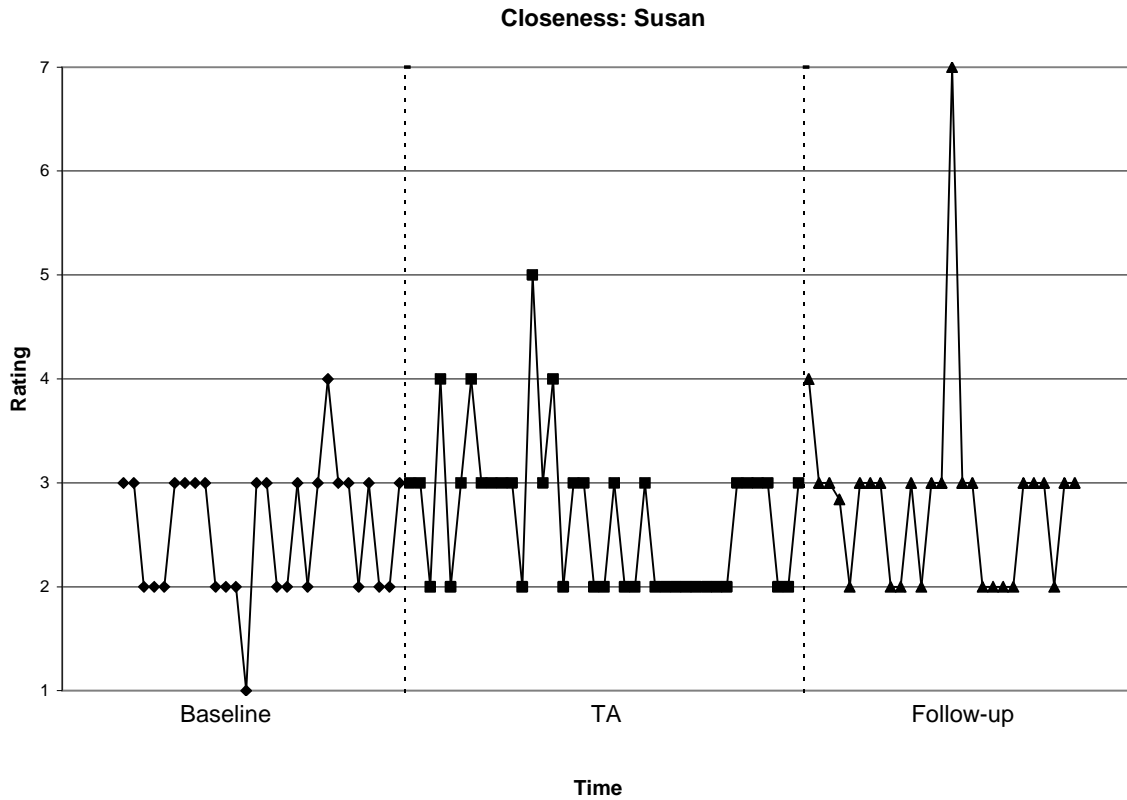
Figure 4.34.



2. How close did you feel to George today? (1=Extremely close, 7=Extremely distant)

- Baseline vs. TA phase: baseline mean (2.54) < TA mean (2.67). $r(\text{Lag } 1) = .09$.
 - No significant level change: $r = .09$; $p = .52$.
 - **Nearly significant slope change: $r = .23$; $p = .08$.**
- Baseline vs. FU phase: baseline mean (2.54) < FU mean (2.85). $r(\text{Lag } 1) = .18$.
 - No significant level change: $r = .19$; $p = .25$.
 - No significant slope change: $r = .09$; $p = .59$.

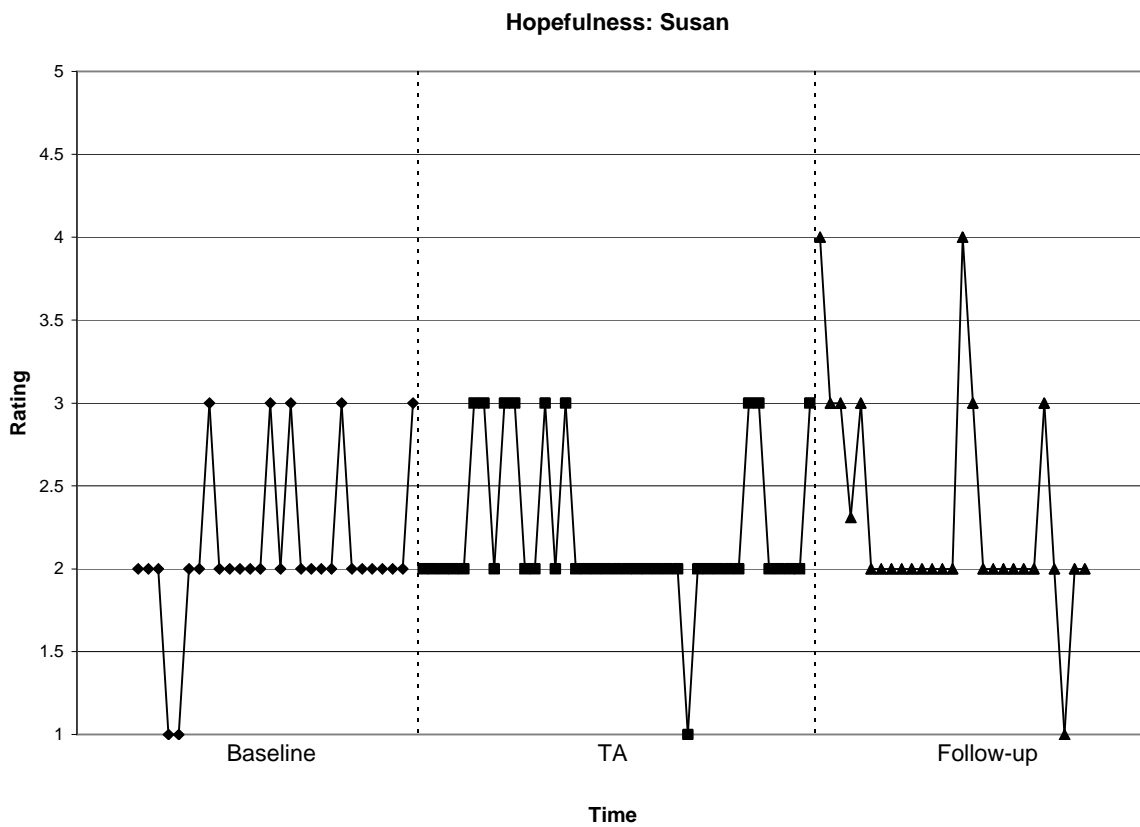
Figure 4.35.



3. How hopeful did you feel about your relationship today? (1=very hopeful, 5=not at all hopeful)

- Baseline vs. TA phase: baseline mean (2.11) < TA mean (2.21) $r(\text{Lag } 1) = .16$.
 - No significant level change: $r = .10$; $p = .50$.
 - No significant slope change: $r = .14$; $p = .34$.
- Baseline vs. FU phase: baseline mean (2.11) < FU mean (2.31) $r(\text{Lag } 1) = .34$.
 - No significant level change: $r = .17$; $p = .37$.
 - **Significant slope change: $r = .37$; $p = .05$.**

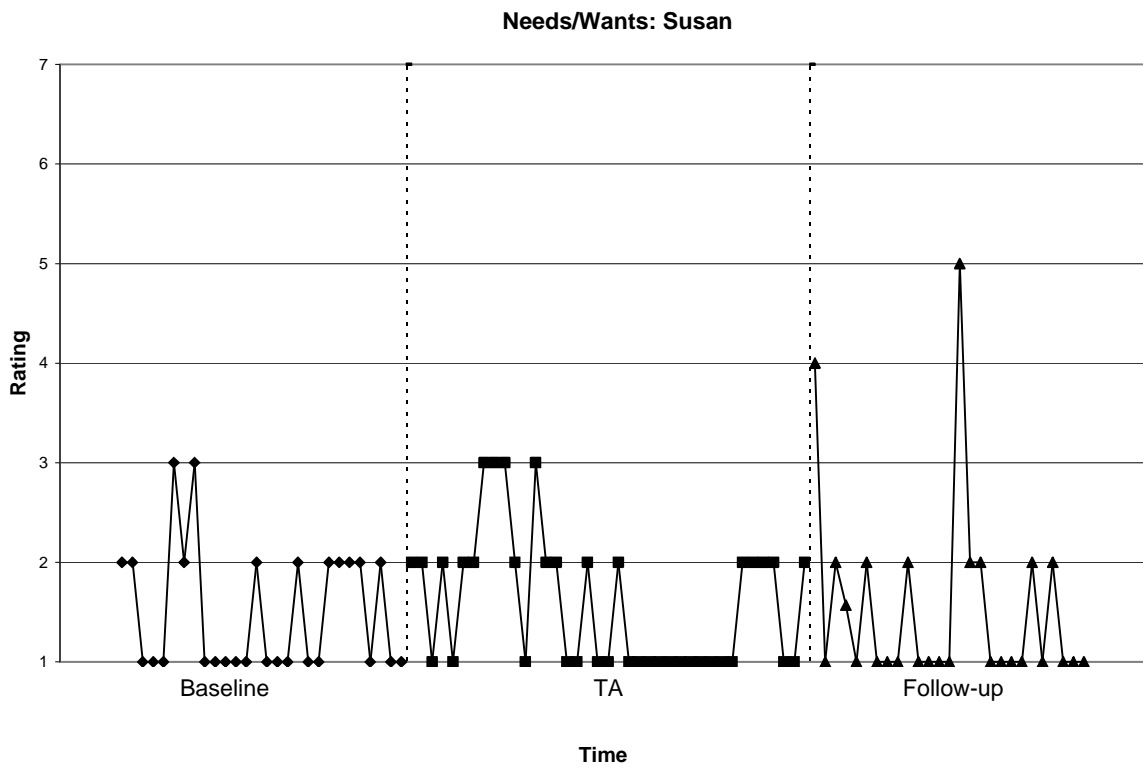
Figure 4.36.



4. On a scale of 1-7 to what degree did you feel that your needs and wants were not as important as your spouse's? (1=not at all, 7=extremely)

- Baseline vs. TA phase: baseline mean (1.50) < FU mean (1.59) $r(\text{Lag } 1) = .30$.
 - No significant level change: $r = .07$; $p = .67$.
 - No significant slope change: $r = .17$; $p = .29$.
- Baseline vs. FU phase: baseline mean (1.50) < FU mean (1.54). $r(\text{Lag } 1) = -.03$.
 - No significant level change: $r = .03$; $p = .88$.
 - No significant slope change: $r = .07$; $p = .60$.

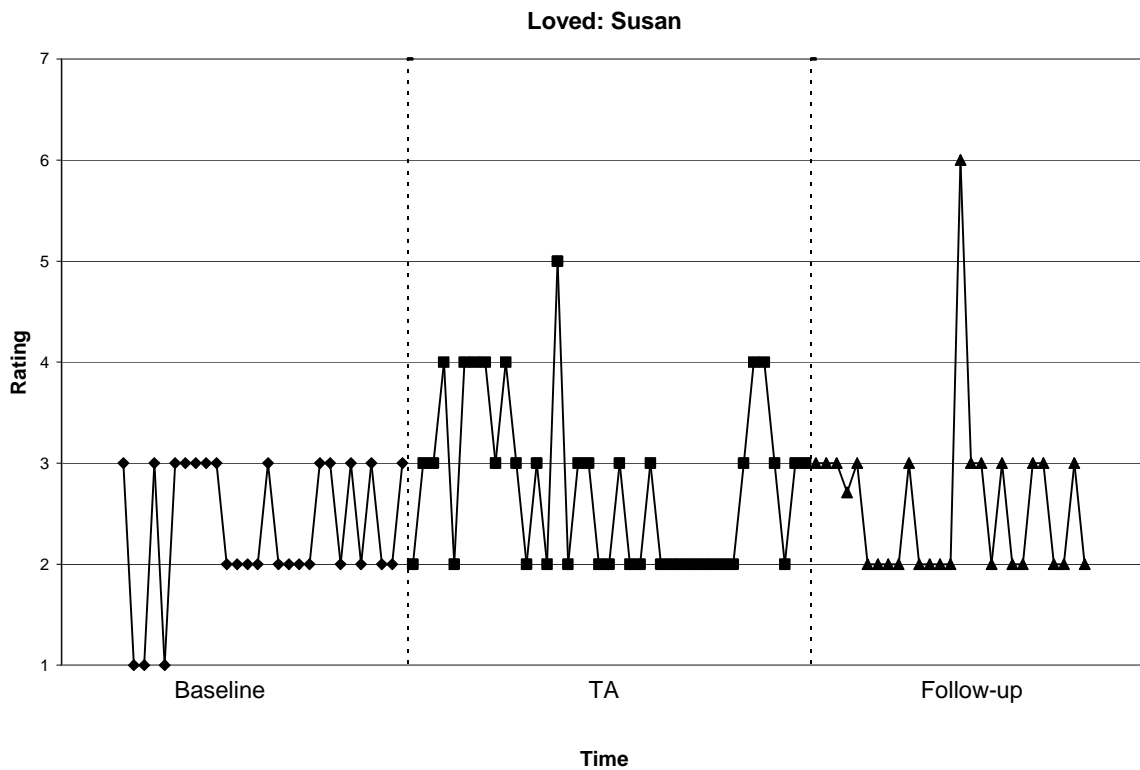
Figure 4.37.



5. On a scale of 1-7, how loved did you feel by George today? (1=extremely loved, 7=not at all loved).

- Baseline vs. TA phase: baseline mean (2.36) < TA mean (2.77). $r(\text{Lag } 1) = .14$.
 - **Nearly significant level change: $r = .26$; $p = .06$.**
 - No significant slope change: $r = .12$; $p = .40$.
 - *Note that the level change indicates she felt less loved during the TA phase.*
- Baseline vs. FU phase: baseline mean (2.36) < FU mean (2.58). $r(\text{Lag } 1) = .02$.
 - No significant level change: $r = .15$; $p = .28$.
 - No significant slope change: $r = .11$; $p = .42$.

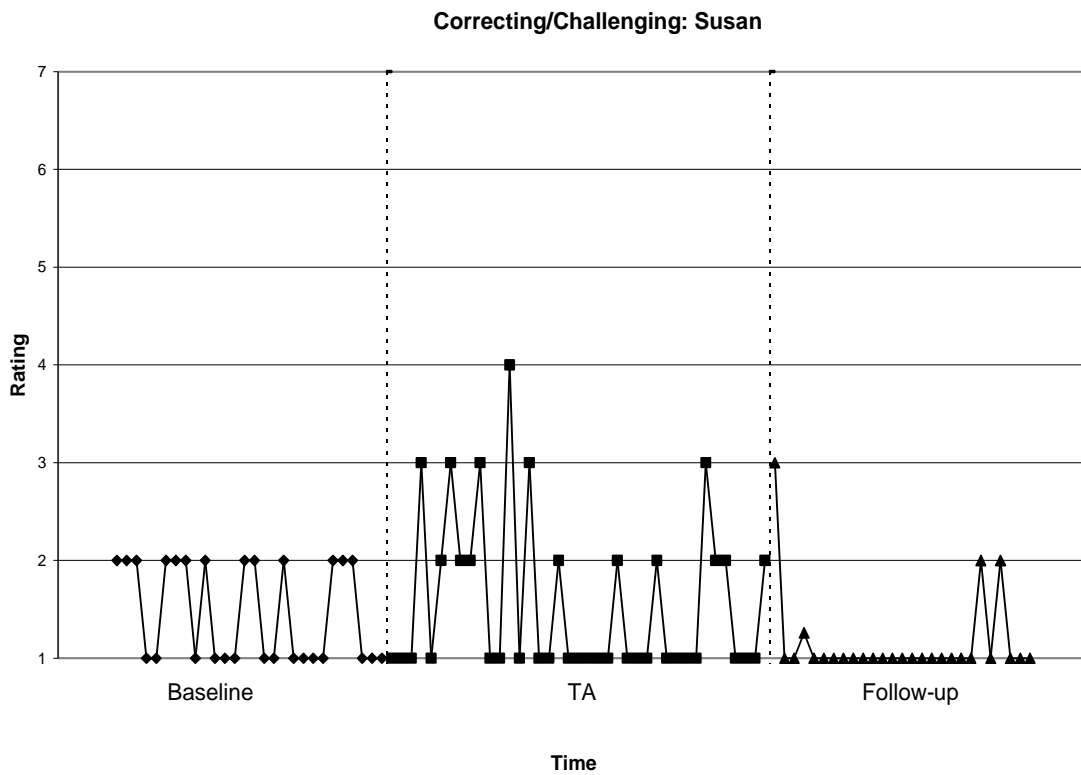
Figure 4.38.



6. On a scale of 1-7, how bothered were you by feelings that George was correcting you or challenging you to defend your opinions today? (1=not at all bothered, 7=extremely bothered)

- Baseline vs. TA phase: baseline mean (1.46) < TA mean (1.56). $r(\text{Lag } 1) = -.02$.
 - No significant level change: $r = .07$; $p = .58$.
 - No significant slope change: $r = .04$; $p = .76$.
- Baseline vs. FU phase: baseline mean (1.46) > FU mean (1.16). $r(\text{Lag } 1) = .13$.
 - **Significant level change: $r = -.31$; $p = .04$.**
 - No significant slope change $r = -.11$; $p = .50$.

Figure 4.39.



Time-series results: Susan

When comparing her baseline ratings to those at follow-up, Susan reported feeling significantly less bothered by feelings that George was correcting her or challenging her to defend her opinions. Although the level of her hopefulness ratings did not improve significantly over time, the slope of these ratings shifted significantly from baseline to follow-up, suggesting her feelings of hopefulness about the relationship were beginning to increase by the end of the follow-up phase; however, on average, she reported slightly lower levels of hopefulness during the follow-up phase, compared to the baseline phase, so it is unclear whether this change in slope represented a clinically significant change. Although her general feelings toward George and her feelings of closeness toward him did not improve significantly across phases, the slope of these ratings shifted somewhat during the TA phase, suggesting she was experiencing an overall increase in closeness and positive feelings over the course of the phase; however, these changes in slope returned to baseline levels during the follow-up phase. Interestingly, Susan reported feeling significantly *less* loved by George during the TA phase when compared to the baseline phase; however, her ratings returned to baseline levels during the follow-up phase.

Clients' perceptions of couples therapy

Visual analysis was used to determine whether the TA was associated with an improvement in George and Susan's perceptions of couples therapy. The means and standard deviations of their Session Evaluation Questionnaire—Client Form (SEQ-C) scores are presented in Table A8 (Appendix K), and the graphs of their weekly SEQ-C ratings are presented in Figures 4.41 and 4.48. Recall again that the SEQ-C was used to assess each partner's feelings about each weekly session with their couples therapist, Linda, throughout all three phases of the study. A visual analysis of George's SEQ-C scores did not reveal any notable differences in his ratings of couples therapy from phase to phase. His phase means do reveal that, on average, he reported higher arousal scores during the follow-up phase. Similarly, Susan also reported higher arousal scores during

the follow-up period, rating these therapy sessions as leaving her feeling somewhat more energetic or excited than sessions in the other two phases. In addition, for Susan, the couples therapy session that was rated as both the “deepest” and the least “smooth” was the second session in the follow-up period.

Figure 4.41.

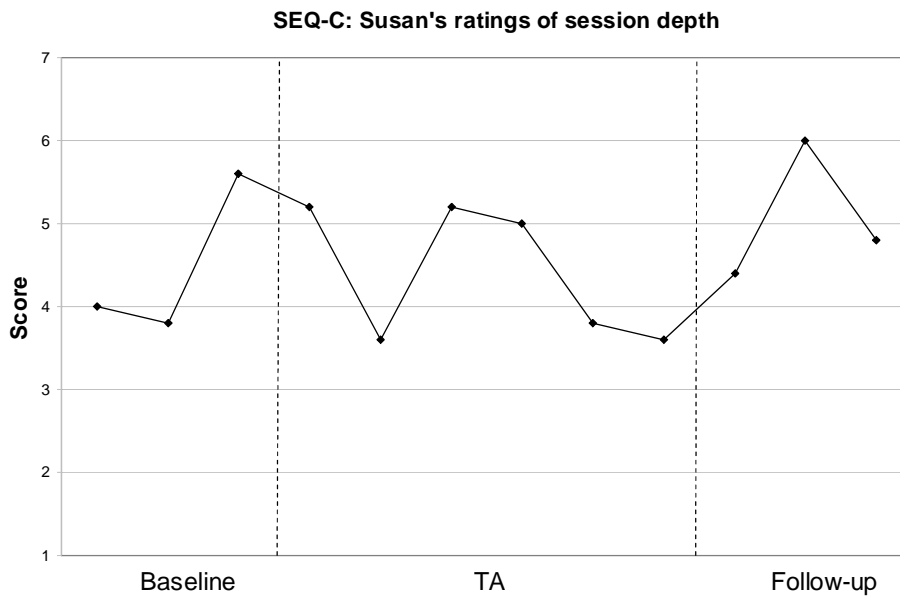


Figure 4.42

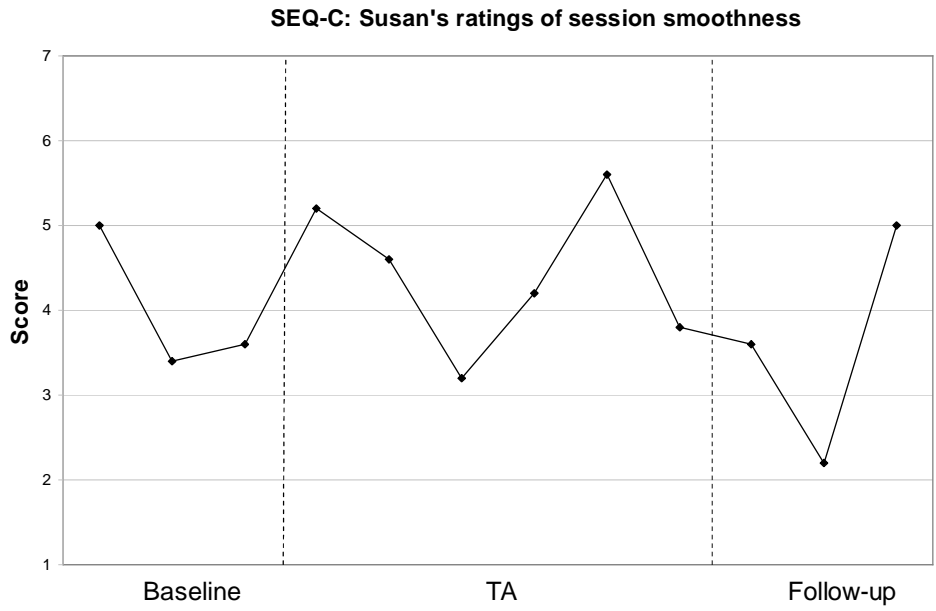


Figure 4.43

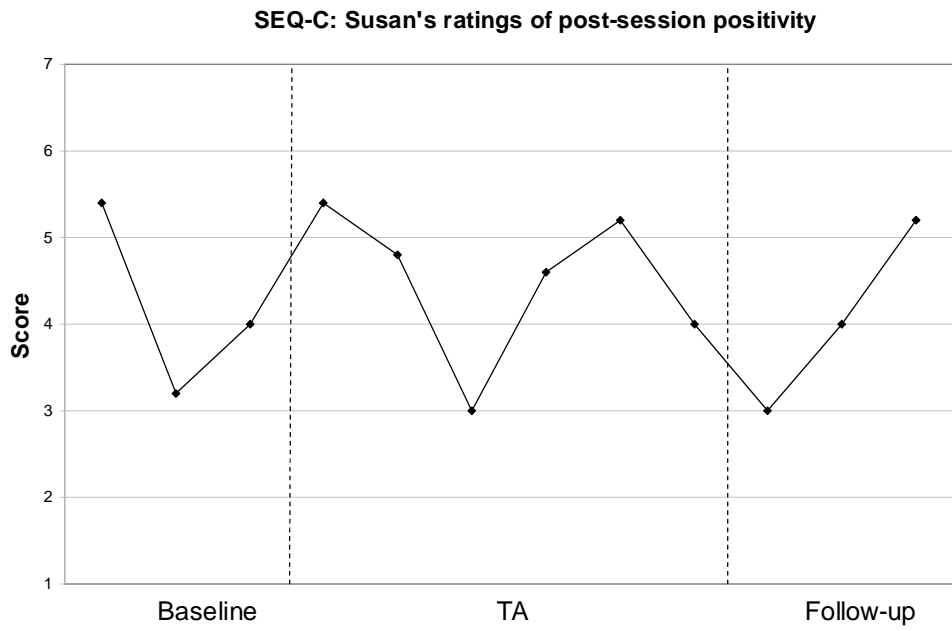


Figure 4.44

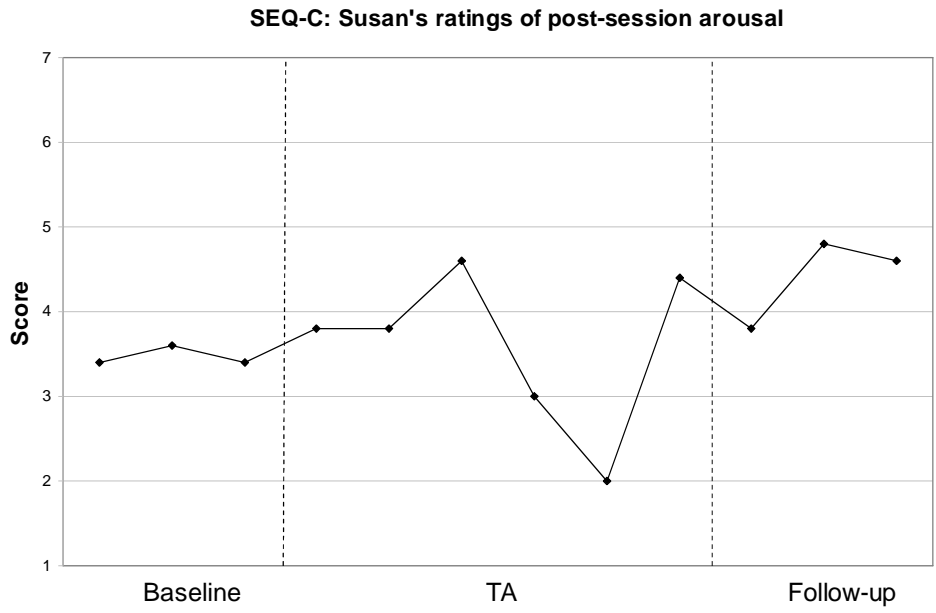


Figure 4.45.

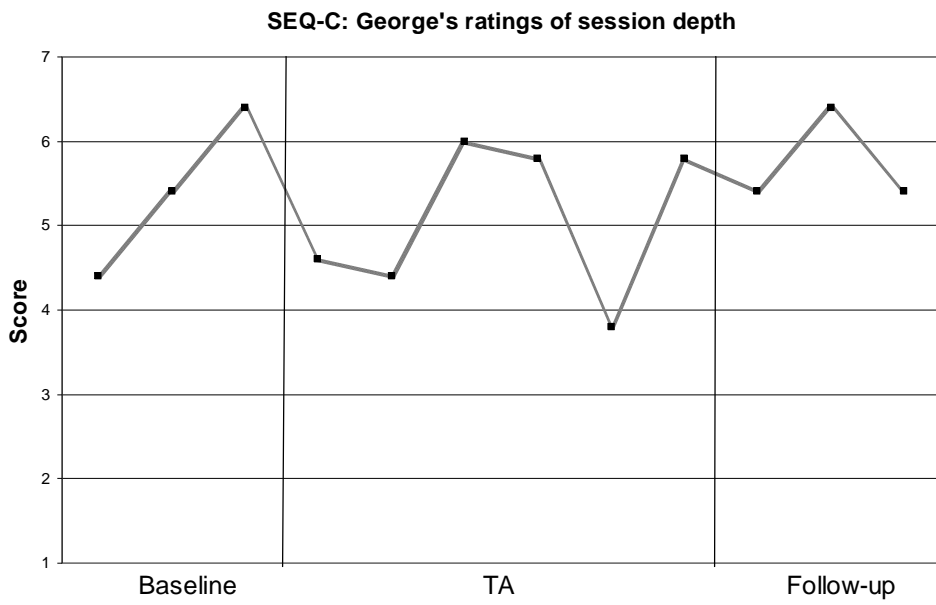


Figure 4.46

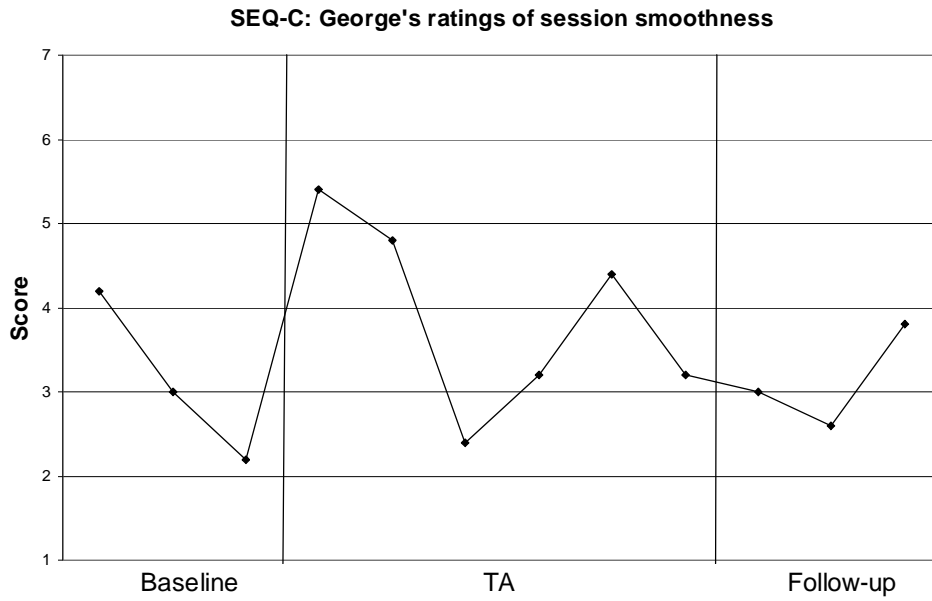


Figure 4.47

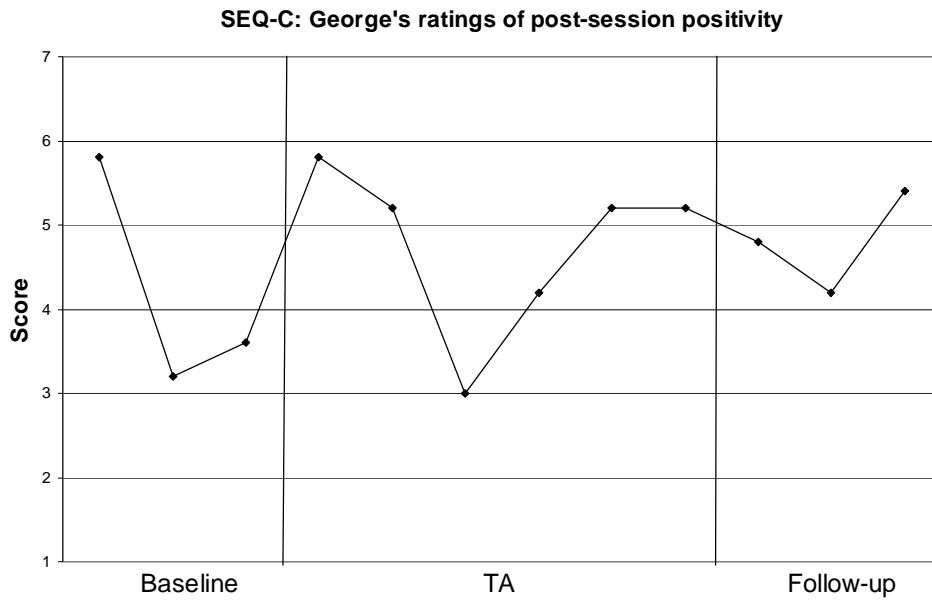
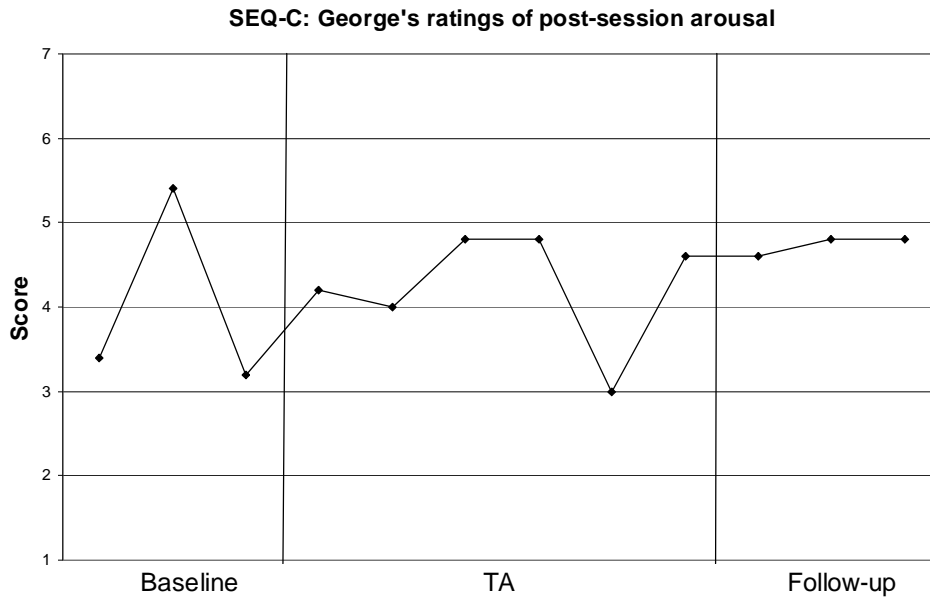


Figure 4.48



Marital satisfaction and psychological symptoms

The couple's BSI, MDS, and DAS ratings are presented in Table A11 (Appendix L), and in Figures 4.49, 4.50, and 4.51. Susan's overall psychological symptoms, as measured by the Brief Symptom Inventory (BSI), decreased by over half a standard deviation, from the clinically distressed range at Time 1 to the non-distressed range and Time 2 and Time 3. George's BSI scores, by contrast, increased by half a standard deviation from Time 1 to Time 3, and all three of his BSI T-scores fell within the clinically distressed range.

Susan's marital satisfaction, as measured by the Dyadic Adjustment Scale (DAS), improved somewhat over the course of the study. According to the cutoff scores that are generally used for the DAS (Sabourin, Valois, & Lussier, 2005), her Time 1 and 2 scores fell in the range that indicates neither the presence nor absence of significant marital distress, while her Time 3 score fell on the cusp of the "absent of significant distress"

range. All three of George’s DAS scores fell in the “absent of significant distress” range, with his highest score occurring at the end of the TA phase (Time 2).

Susan’s scores on the Marital Disaffection Scale (MDS) decreased over the course of the study, indicating she felt lower levels of “disaffection” or “loss of love” for George by the end of the study. According to Kayser’s (1996) categories, Susan’s MDS scores moved from the “above average” range of disaffection to the “average” range. George’s MDS scores, by contrast did not change substantially from baseline to follow-up, and all three of his scores fell in the “below average” range of disaffection.

Figure 4.49

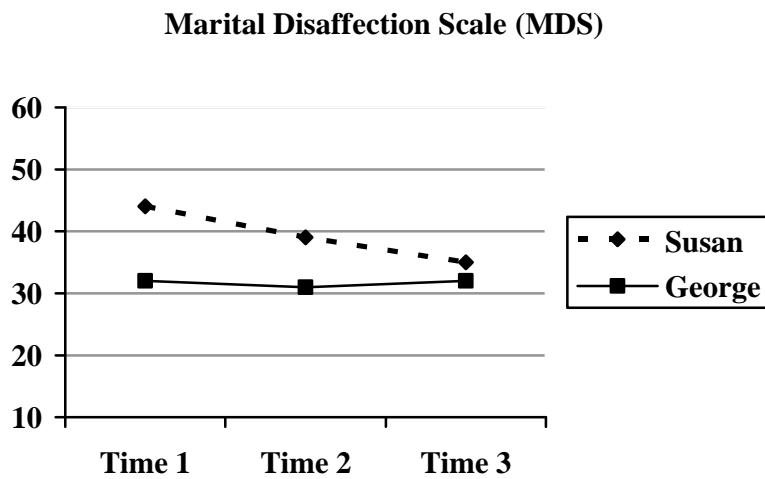


Figure 4.50.

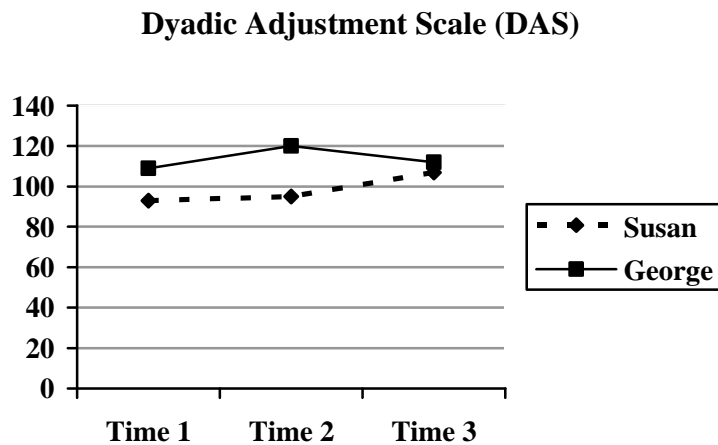
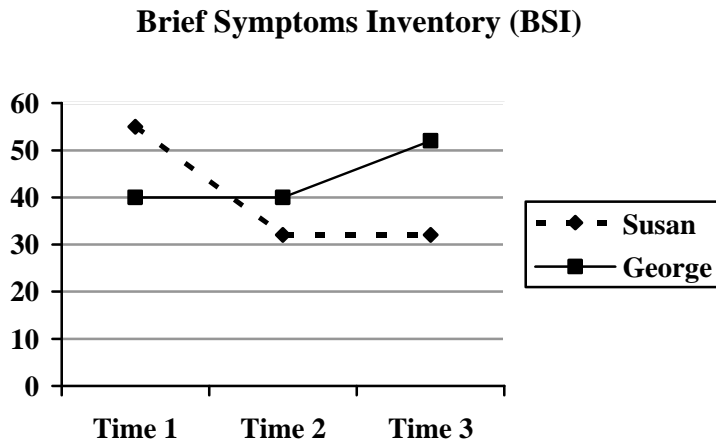


Figure 4.51.



Couple's perceptions of Therapeutic Assessment

Susan and George's T-scores on each of the subscales of the Assessment Questionnaire (AQ-2) are presented in Table 4.4. On most of the subscales, both partners' scores fell within half a standard deviation of the mean, indicating their satisfaction with the TA was average when compared to most clients receiving psychological assessments at the Center for Therapeutic Assessment. The one exception was Susan's T-score of 57.6 on the "positive relationship with the examiner" subscale, which was more than a half of a standard deviation above the mean.

Table 4.4. *AQ-2 Results (Mean Ratings and T-scores)*

	New self-awareness/self-understanding	Positive accurate mirroring	Positive relationship with the examiner	Negative feelings about the assessment	Total satisfaction
Susan	Mean=3.69 T=49.0	Mean=3.33 T=48.9	Mean=4.58 T=57.6	Mean=2.00 T=51.7	Mean=3.85 T=52.3
George	Mean=3.85 T=50.4	Mean=3.17 T=47.8	Mean=3.58 T=46.5	Mean=2.33 T=53.5	Mean=3.52 T=49.3

Note: Participants rated each item on a 1-5 Likert-type scale, with high scores indicating higher satisfaction, with the exception of the "negative feelings about the assessment" scale, on which higher ratings indicated more negative feelings.

Couple's qualitative feedback about the TA

Susan:

Susan provided a mix of positive and negative feedback about her experience with the TA. When asked how well the assessment met her expectations, she replied, "I thought the individual assessments were very good and met my expectations. I wanted more time on how these issues show in our couples relationship, but this may happen in our ongoing therapy." When asked about the most valuable parts of the assessment, she replied, "The individual assessments were good although the results were not that surprising-George is detached from his emotions, I am over sensitized to the emotional state of others - they did crystalize these issues for us." When asked about the "least valuable" parts, she replied, "I don't think the answers to our questions were as valuable as I had hoped." Regarding how the assessment affected her relationship with George, she stated that "the day of the assessment it brought us together, but we had a big fight the very next morning. I really felt hurt when George tuned me out and just didn't want to 'put up with it.'" When asked how the assessment affected her experience in couples therapy, she replied, "I think it has enhanced it. I don't think it has altered the direction of the therapy. I think it would have been even more helpful earlier in our therapy." When asked for suggestions to improve the way assessments are done at the Center for Therapeutic Assessment, she replied, "I think the assessment is not described correctly. To me it is about the foundation of problems in a relationship, not about the therapy. I mean, Linda wasn't missing anything so if it's an assessment of her she gets an A." She reported she would tell a friend that "I think the assessment could be very helpful in setting out the major issues in a relationship." She also commented, "I think that the regular therapist should have individual sessions already scheduled with each participant shortly after the assessment is done."

George:

George's feedback about the TA was generally positive. When asked how well the assessment met his expectations, George replied, "Well. In fact there was more there

in terms of specific answers and advice than I was expecting.” He listed the following aspects of the assessment as being “most valuable:” “1. Confirmation that I am not connected to my feelings. 2. A better look at what my actions/reactions do to others. 3. Some more insight into what I am really responding to when I react/act. 4. More insight into what Susan is responding to when she reacts/acts.” When asked about the “least valuable” parts of the assessment, he replied, “There wasn’t a ‘bad’ part. It all seemed at the least interesting and probably useful -- Knowledge is power. (Francis Bacon).” When asked how he thought the assessment affected his relationship with Susan, he replied,

It's a little early to tell. The revelation that Susan has long-running depression will surely have an impact, and the suggested plan for a sex life has me pretty hopeful. A lot of the other data, if we pay heed, should make our life together --I don't know, more pleasant?, richer?, easier?--all of these, maybe. The very act of doing of the assessment together I think had a good effect.

Regarding how the assessment affected his experience in couples therapy, he stated:

1. It confirmed several of the hypotheses under which we were working with Linda. 2. The data on Susan and Depression has got to be really helpful. 3. The act of working together on the assessment was a good thing.

When asked about suggestions to improve the way assessments are done at the Center for Therapeutic Assessment, he replied,

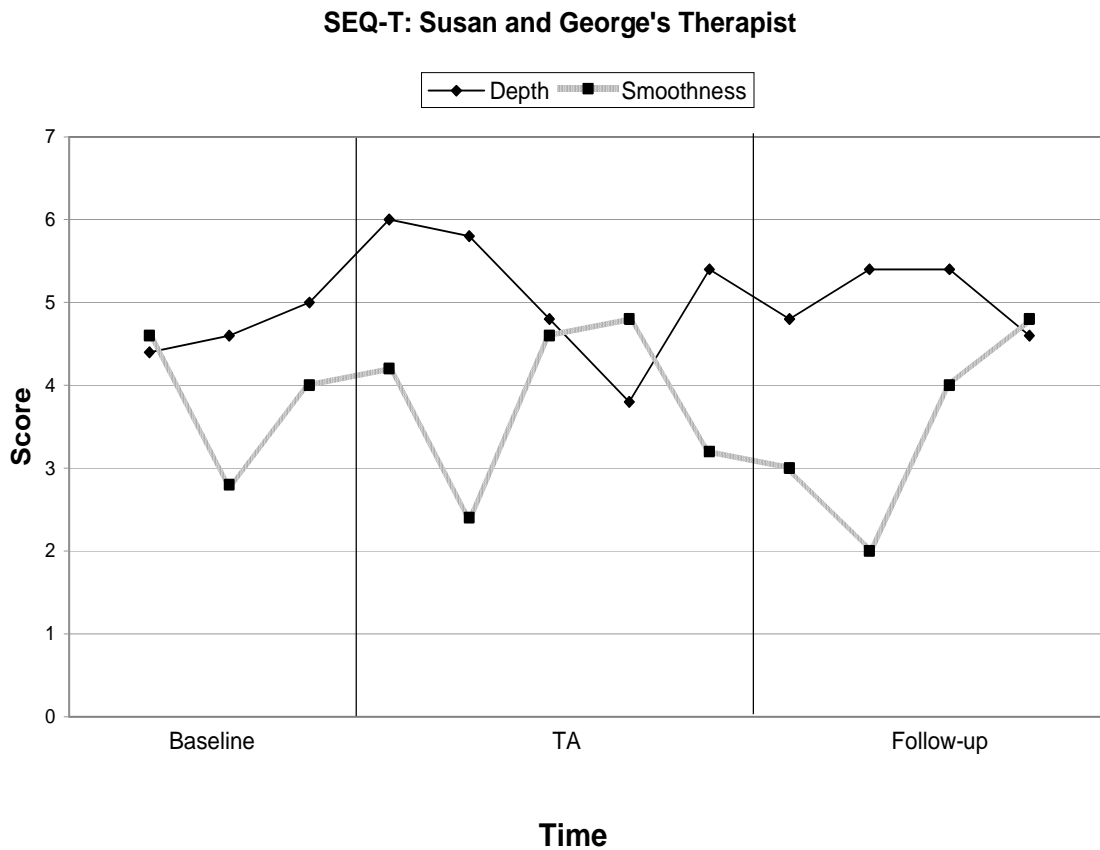
I thought the whole experience was good. So much time was spent making us feel at ease, and describing us in favorable terms (I am a Stout Survivor- have there been any Frail Finishers?) that I started to worry about whether the truth might suffer from the efforts to do no harm, but in the end Dr. S handled things so well that I felt a little bad for having had any doubt. No changes.

He stated he would tell a friend considering an assessment that, “at a minimum you will find out interesting things, and probably things w/ the potential, if heeded, for really good life changes. Do it.”

Therapist's ratings of couples therapy—Susan and George

To ascertain Susan and George's therapist's perceptions of couples therapy before, during, and after the TA, her ratings on the Session Evaluation Questionnaire—Therapist Form (SEQ-T) were examined. The means and standard deviations of her SEQ-T scores are presented in Table A14 (Appendix M) and a chart of her weekly SEQ-T ratings is presented in Figure 4.52. The therapist's ratings of the couples therapy sessions did not appear to change dramatically from phase to phase. It may be of note that the two sessions she rated as "deepest" occurred at the beginning of the TA phase. However, these scores were not substantially different from depth scores in other phases, and it is unclear from the data whether the brief increase in depth was in any way related to the TA.

Figure 4.52.



Therapist's qualitative feedback—Susan and George

Susan and George's therapist reported having largely positive impressions of the TA and its effect on her therapy with the couple. She reported having the following "hopes and expectations" when she referred the couple for a psychological assessment:

My hope was for me and them to get more information to explain their blocks to physical and emotional intimacy. I hoped the couple could gain some conscious "knowing" of their underlying issues and develop compassion for self and other.

At the conclusion of the assessment, she reported, "The assessment validated many ideas I had and helped the couple to understand at a deeper level. It feels like it strengthened the therapy process—but didn't necessarily make a quantitative leap." When asked about the "most useful" part of the assessment, she replied:

Susan clearly felt that the daily questions were most helpful—she had to acknowledge the "data" did not meet the tapes in her head. George said it was the "whole" thing. For me, the assessment facilitated addressing certain issues and "meta-processing" our work together.

She reported the "least helpful" part of the assessment was "not being able to give them more information on the MMPI's and how their profiles intersected." When asked what would have made the assessment more useful, she commented, "Of course the Rorschach would have helped—but with the shortened version—more than one feedback session or spread them out." She said she would tell a colleague who was considering referring a client for this type of assessment that "it can be extremely helpful and determine if a couple and therapeutic relationship can manage the 'disruption.'"

Discussion: George and Susan

An analysis of their daily ratings indicated George and Susan were able to make small improvements on some of their presenting problems. For example, George's ratings of feeling "loved" by Susan improved significantly from baseline to follow-up, and Susan reported feeling less bothered by George correcting and challenging her by the follow-up phase. Many of their presenting problems, however, showed no significant progress by

the end of the follow-up phase. For example, one of their original treatment goals, improving their sex life, was not accomplished. It may be that this problem was too longstanding and entrenched in each partner's own personality dynamics to be resolved by a short-term intervention. Although it is possible that the TA opened up insights and channels of communication that may have eventually led to improved sexual intimacy, it is impossible to know from the short, four-week follow-up period whether the couple was ever able to make headway in this area of their relationship. In addition, as with Mary and Ray, it is difficult to know whether statistically significant changes on the daily measures represented *clinically* significant changes in the couple's daily life together.

Time-series results also revealed that, on a few variables, things between George and Susan actually worsened during the TA phase, suggesting that the intervention may have "stirred up" difficult emotions and interpersonal dynamics for the couple. It is unclear, however, whether this process was a helpful or necessary piece of the intervention. In any case, the fact that these ratings returned to baseline levels during the follow-up phase indicates the couple's relationship was able to withstand, if not grow from, any disruption caused by the TA.

Judging from the SEQ ratings, is unclear whether the TA had any substantial impact on George and Susan's therapy with Linda. Although both partners reported slightly higher levels of arousal during couples therapy sessions in the follow-up period, it is difficult to determine whether this was due to the TA or other factors. For the most part, George, Susan and Linda all seemed to view the therapy as having ups and downs throughout the course of the study, and these fluctuations did not seem to follow any discernable pattern. Indeed, Linda's observation that the TA was helpful, but "didn't necessarily make a quantitative leap" appears accurate. This lack of change may be the result of the couple having been in therapy with Linda (who is regarded in the community as a "premier" couples therapist) for three years; both partners had reported feeling very satisfied with their therapy at the outset of the study, and it may be that the couple and their therapist had settled into a rhythm together that would have been very difficult (and perhaps unnecessary) to change.

One of the most interesting findings for Susan and George was the change in their BSI scores across the course of the study. While Susan's scores moved from the distressed to non-distressed range from baseline to follow-up, George's psychological symptoms, as measured by this instrument, increased over time, with all three of his scores falling in the clinically distressed range. It is notable that one partner became more symptomatic as the other partner became less so. This is similar to the pattern seen in Mary and Ray's BSI scores, and again raises questions about whether the TA led to a shift in the emotional equilibrium of the relationship. Although one can only speculate, it may be that George came to acknowledge or experience more symptoms as Susan began to relinquish her identity as the "sick" member of the couple. Similarly, Susan's scores on the two measures of marital satisfaction (MDS and DAS) showed small improvements, while George's scores essentially remained the same. It is important to note however, that George's MDS and DAS scores were in the non-distressed ranges at Time 1, so there was little room for improvement at Times 2 and 3. These findings, in which Susan improved and George did not, are somewhat curious given that George provided more positive qualitative feedback about the TA than Susan. The fact that the couple completed the qualitative feedback questionnaire immediately after the TA may at least partially account for this. We see from his daily ratings that George felt the most hopeful at the end of the TA phase, but that this sense of hopefulness was not sustained during the follow-up phase. It is possible that the sense of optimism he felt right after the TA was short-lived; or, it may be that the results of the TA were simply mixed for both partners: it was helpful in some ways, and less-helpful (perhaps even painful) in others.

Couple #3: Elaine and Tom

Elaine and Tom were in their early 30s and had been married for just under two years at the outset of the study. Tom, a white man, had a Bachelor's degree and worked as a consultant. Elaine, a white woman, had a graduate degree and worked in the technology industry. They had been seeing their therapist, Carla, for just under a year. They originally presented for couples therapy because they were arguing frequently and needed help with communication. They reported they would "argue and get nowhere . . . we were talking different languages." Both partners agreed that therapy had given them "an outlet" for their emotions; they said Carla pushed Elaine to talk about "uncomfortable" topics, while she "called Tom out" on his tendency to be over-analytical. At the time of the study, they noted that many of their marital problems had to do with their difficulties coping with external stressors. Both reported experiencing significant problems with their own families of origin. Tom often felt as though Elaine's family members were "bullying" her, leaving Tom feeling protective and frustrated. Similarly, Tom's tendency to get caught up in his parents' and siblings' problems often left Elaine feeling angry and left out. They remarked that these outside stressors had become so consuming that they felt they were losing their "identity as a couple." Both commented that they would like to be able to focus more on their marriage, rather than external factors. They also remarked that they felt they were perhaps living too much in the present without planning for the future as a couple (e.g., around finances, having children). In addition, Elaine reported experiencing several symptoms of anxiety, particularly in relation to problems with her parents and siblings. Tom reported experiencing several symptoms of tension and irritability.

After the Time 1 interviews with Tom and Elaine, the researcher used information from the interview, as well as the couple's responses to the DAS, MDS, and BSI, to develop a set of questions to which each participant would respond on a daily basis throughout the study. As with the other two couples, the researcher sent drafts of the questions to both partners and invited them to edit or amend the questions as they saw fit. Each of these questions is presented along with the time-series results.

Integrity rating of TA sessions

Treatment integrity ratings indicated Tom and Elaine’s assessor conducted the assessment according to the principles of Therapeutic Assessment as described by Finn (2007) (see Table 4.5).

Table 4.5. *Treatment integrity ratings—Tom and Elaine*

Initial session with couple	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. The assessor sought assessment questions from both clients.					X
2. The assessor asked the clients about their “best guess” answers to at least one of these questions.					X
3. The assessor let the clients know that he/she would be asking for their input/collaboration throughout the assessment process.					X
4. The assessor talked to the clients about the tests they would be taking.					X
5. The assessor demonstrated empathy.					X
6. The assessor demonstrated good listening skills.					X
7. The assessor encouraged the clients to ask questions about the assessment.					X
8. If a client asks a non-systemic question about his/her partner, the assessor helped to rephrase this as a systemic question.					n/a
Consensus Rorschach /Assessment Intervention session	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. During the Consensus Rorschach task, the assessor asked the couple to come up with responses that they could both see and agree on.					X
2. After administering the Consensus Rorschach, the assessor asked the clients to talk about their experience during the task (e.g., thoughts, feelings, observations).					X

Table 4.5, cont.

3. The assessor attempted to make connections between what happened during the Consensus Rorschach task and what happens in other areas of the couple's life together.					X
4. The assessor asked the couple to try out a new approach to the Consensus Rorschach (using a card that was not used during the first administration).					X
5. The assessor demonstrated empathy.					X
6. The assessor demonstrated good listening skills.					X
7. The assessor made systemic interpretations about how each person contributes to the patterns revealed.				X	
Summary/Discussion session	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. The assessor appeared to begin by presenting Level 1 findings, followed by Level 2 and (if appropriate) Level 3 findings.				X	
2. The assessor encouraged the clients to revise or amend the findings.					X
3. The assessor demonstrated empathy.					X
4. The assessor demonstrated good listening skills.					X
5. The assessor seemed to give equal emphasis to each partner's contribution to the relationship struggles.				X	
6. The assessor asked clients to give real life examples of assessment findings that were discussed.					X

Results: Elaine and Tom

Qualitative information from daily surveys

The following prompt was included on each of the couple's daily surveys: "If anything happened today that you think affected your feelings about your relationship in any way, feel free to mention it in the box below." Tom did not use the comments section for this purpose at any point during the study. Elaine did include a few comments with her daily surveys, and those comments are summarized here.

Most of Elaine's comments were related to stressors involving her and Tom's family members (e.g., her sister's pregnancy, Tom's father's health). For example, midway through the baseline phase, she wrote, "Work, and my sister's troubled pregnancy both have me a bit anxious and focused elsewhere." She also noted that both she and Tom had job interviews during the second-to-last week of the TA phase. She did not write any comments during the follow-up phase, except to note that she and Tom took a vacation together during the final week of the phase.

Time-Series Phase Effect Analyses: Tom

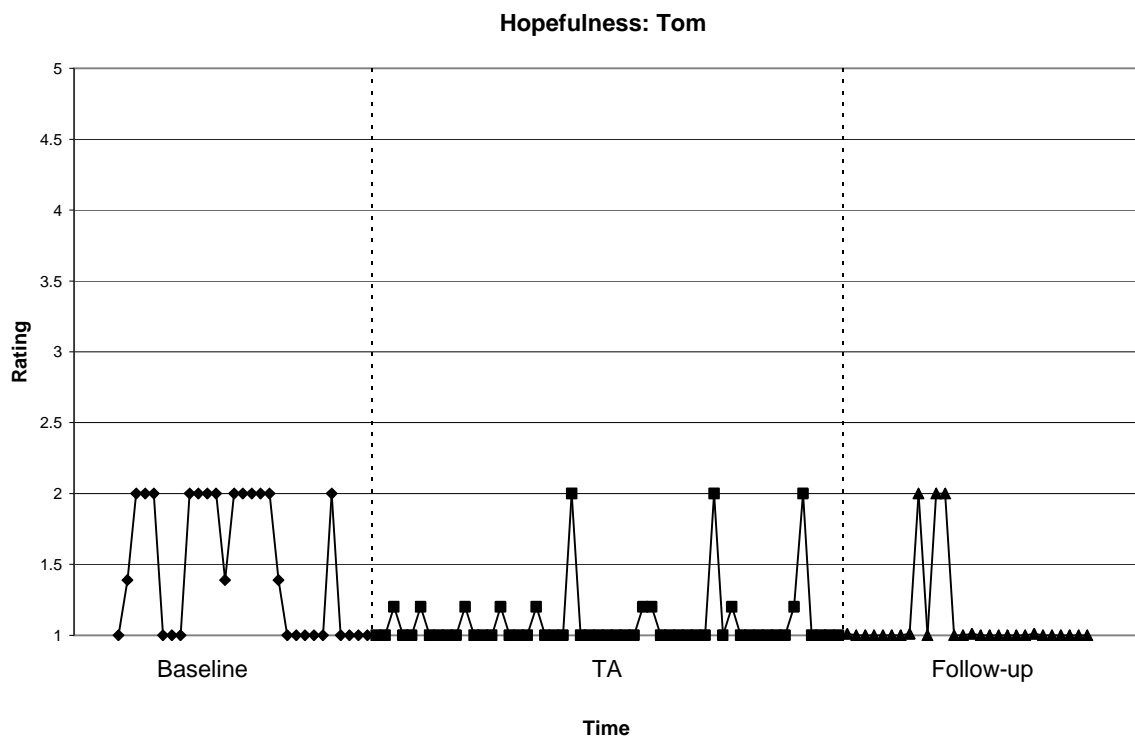
The means, standard deviations, OLS slopes, and number of data points for each of Tom's daily questions are presented in Table A5 (Appendix J). For each daily question, phase effect analyses were run to determine whether there was a significant change in the slope and level of the scores from the baseline to the TA phase, and from the baseline to the follow-up (FU) phase. The estimation maximization procedure (EM) was used to estimate missing data values. Data were missing for 14.5% of Tom's daily questions. Little's MCAR test yielded a chi-square of .184 and p-value of .668, indicating the MCAR assumption was met.

Results of the phase effect analyses, along with line graphs of Tom's ratings on each of the daily questions are presented below. Significant ($p < .05$) and nearly significant ($p < .10$) results are listed in **bold**.

3. How hopeful did you feel about your relationship today? (1=very hopeful, 5=not at all hopeful)

- Baseline vs. TA phase: baseline mean (1.49) > TA mean (1.09). $r(\text{Lag } 1) = .52$.
 - **Significant level change: $r = -.49$; $p = .003$.**
 - No significant slope change: $r = .06$; $p = .76$.
- Baseline vs. FU phase: baseline mean (1.49) > FU mean (1.11). $r(\text{Lag } 1) = .57$.
 - **Nearly significant level change: $r = -.43$; $p = .07$.**
 - No significant slope change: $r = -.15$; $p = .55$.

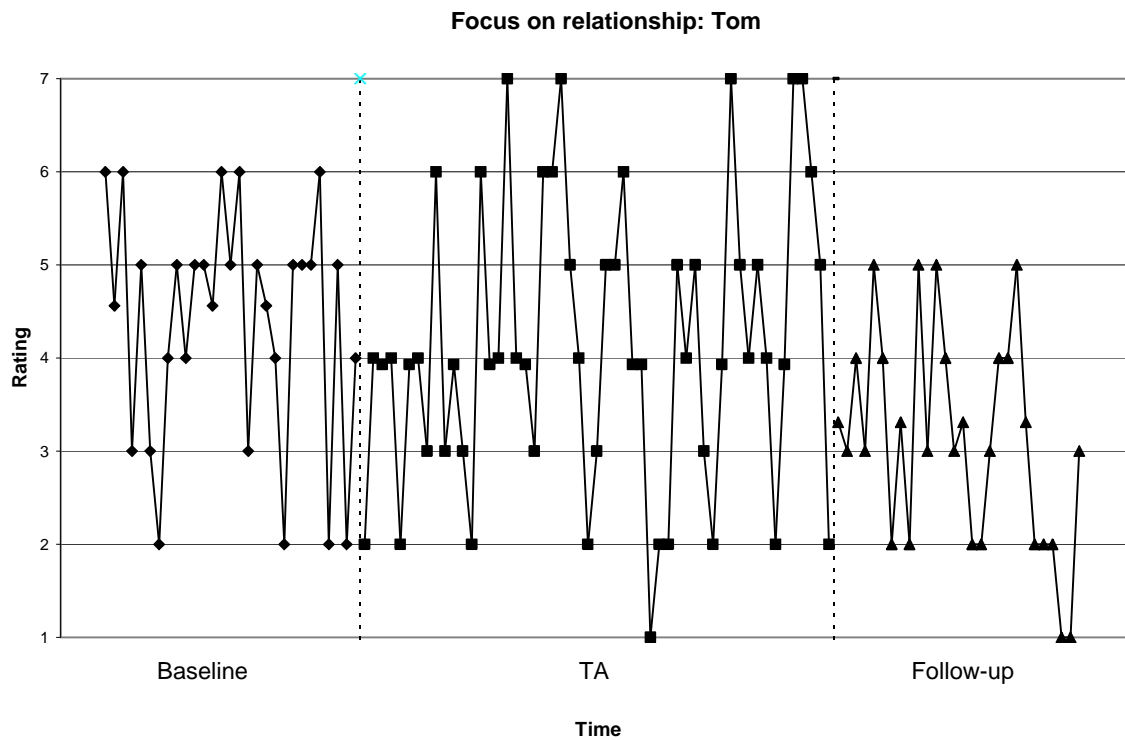
Figure 4.55.



4. To what degree did you feel your attention was focused on your relationship with Elaine, versus other things happening in your life (e.g., things going on at work or with other family members or in-laws)? (1=focused almost exclusively on the relationship, 7=almost exclusively focused on other things)

- Baseline vs. TA phase: baseline mean (4.37) > TA mean (4.14). $r(\text{Lag } 1) = .18$.
 - No significant level change: $r = -.08$; $p = .55$.
 - No significant slope change: $r = -.11$; $p = .42$.
- Baseline vs. FU phase: baseline mean (4.37) > FU mean (3.12). $r(\text{Lag } 1) = .27$.
 - **Significant level change: $r = -.46$; $p = .01$.**
 - No significant slope change: $r = .07$; $p = .69$.

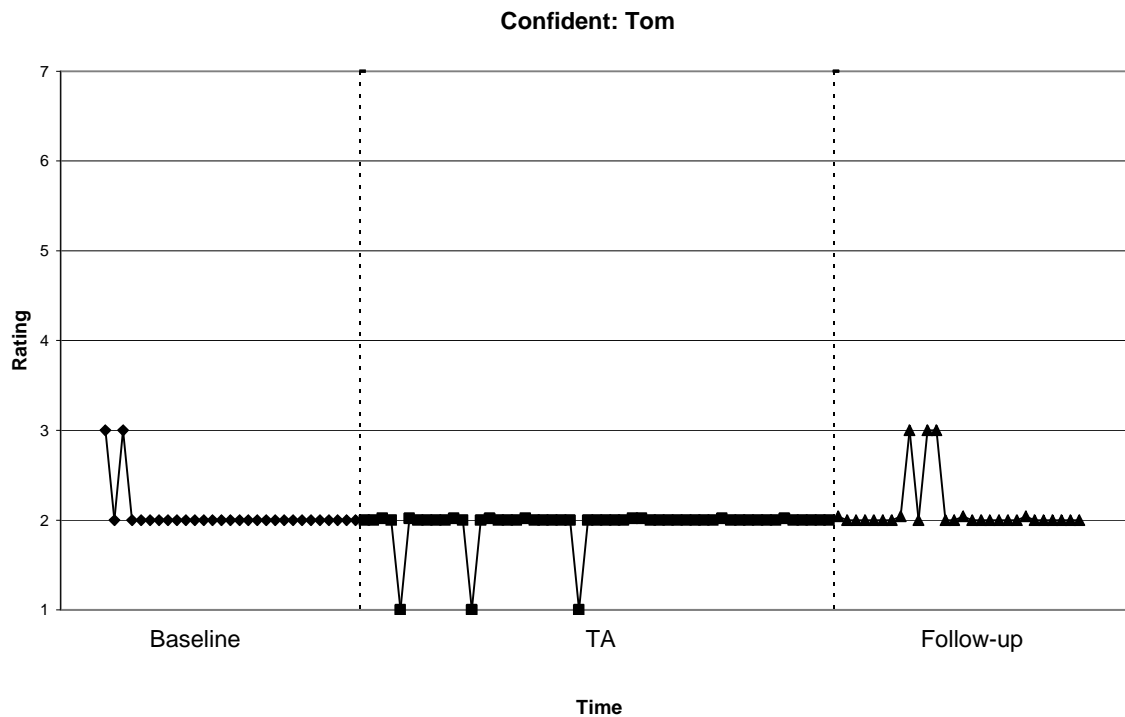
Figure 4.56.



5. On a scale of 1-7, to what degree did you feel confident that your relationship is based on a solid foundation? (1=extremely confident, 7=not at all confident)

- Baseline vs. TA phase: baseline mean (2.07) > TA mean (1.95). $r(\text{Lag } 1) = -.01$.
 - **Significant level change: $r = -.24$; $p = .03$.**
 - No significant slope change: $r = -.12$; $p = .27$.
- Baseline vs. FU phase: baseline mean (2.07) < FU mean (2.11). $r(\text{Lag } 1) = .15$.
 - No significant level change: $r = .08$; $p = .61$.
 - No significant slope change: $r = -.10$; $p = .52$.

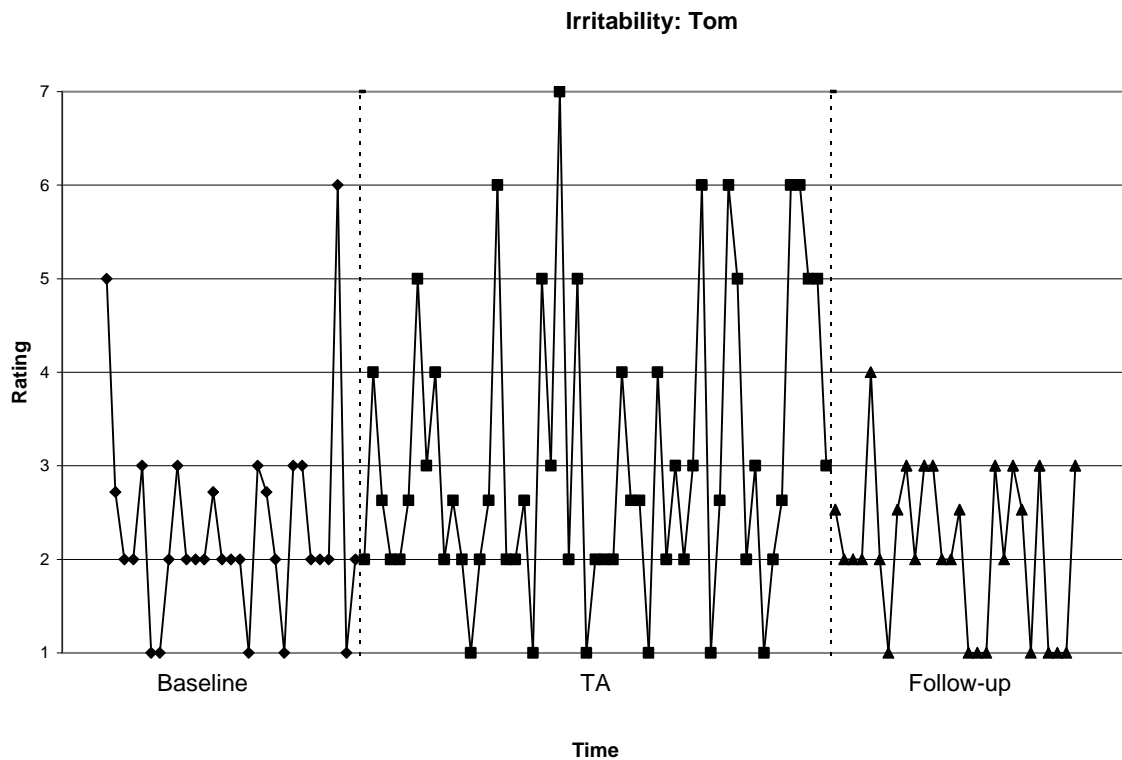
Figure 4.57.



6. On a scale of 1-7, how tense or irritable did you feel today? (1=not at all tense or irritable, 7= extremely tense or irritable)

- Baseline vs. TA phase: baseline mean (2.32) < TA mean (3.07). $r(\text{Lag } 1) = -.07$
 - **Significant level change: $r = .25$; $p = .03$.**
 - **Significant slope change: $r = -.26$; $p = .03$.**
 - Note that both the slope and level changes indicate an **increase** in irritability during the TA phase.
- Baseline vs. FU phase: baseline mean (2.32) > FU mean (2.11). $r(\text{Lag } 1) = .08$.
 - No significant level change: $r = -.11$; $p = .42$.
 - No significant slope change: $r = .09$; $p = .47$.

Figure 4.58.



Summary of time-series results: Tom

When comparing his ratings at baseline to those at both the TA and follow-up phases, Tom's feelings of closeness to Elaine increased significantly. He also reported feeling significantly more hopeful about his relationship with Elaine during the TA and follow-up phases when compared to the baseline phase; this difference was statistically significant for the TA phase and nearly significant for the follow-up phase. In addition, he reported feeling significantly more focused on his relationship with Elaine at follow-up, compared to baseline. He also reported feeling significantly more confident in his relationship with Elaine during the TA phase; however, his ratings on this variable returned to baseline levels during the follow-up phase. Interestingly, he reported feeling significantly more irritable during the TA phase, compared to the baseline phase, and the significant change in slope suggested a trend toward greater irritability over the course of the TA phase; however, his irritability ratings and their slope returned to baseline levels during the follow-up phase.

Time-Series Phase Effect Analyses: Elaine

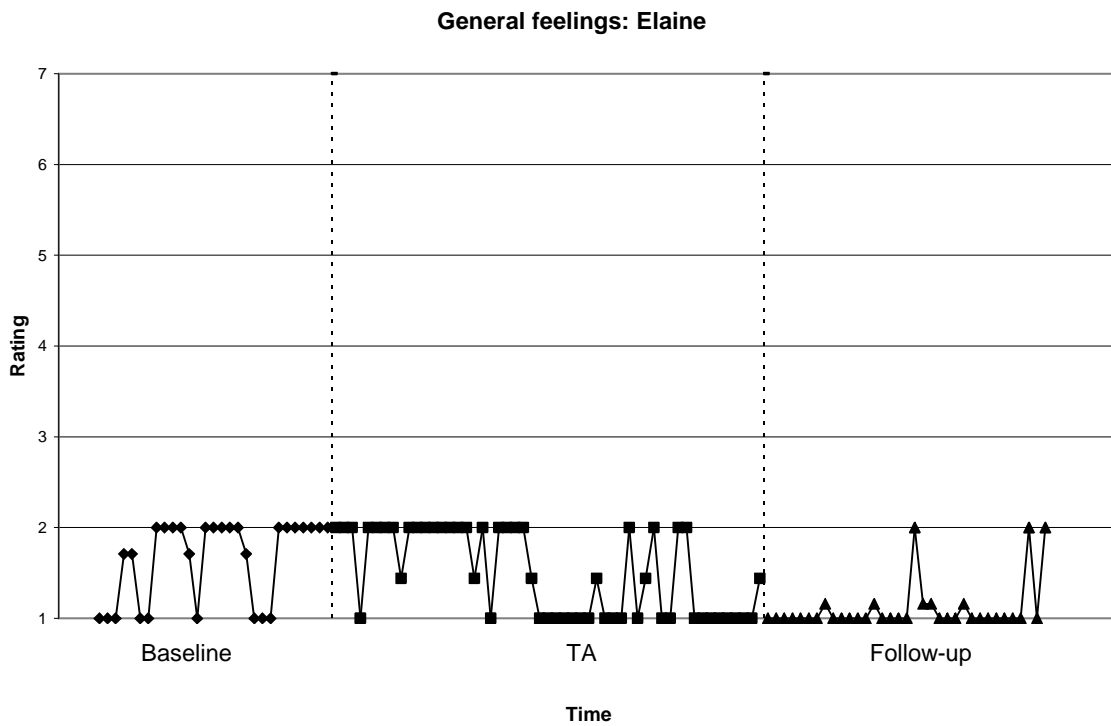
The means, standard deviations, OLS slopes, and number of data points for each of Elaine's daily questions are presented in Table A6 (Appendix J). For each daily question, phase effect analyses were run to determine whether there was a significant change in the slope and level of the scores from the baseline to the TA phase, and from the baseline to the follow-up (FU) phase. The estimation maximization procedure (EM) was used to estimate missing data values. Elaine Data were missing for 12.8% of Elaine's daily questions. Little's MCAR test yielded a chi-square of .007 and p-value of .931, indicating the MCAR assumption was met.

Results of the phase effect analyses, along with line graphs of Elaine's ratings on each of the daily questions are presented below. Significant ($p < .05$) and nearly significant ($p < .10$) results are listed in **bold**.

1. Please rate your general feelings toward Tom today (1=Extremely positive, 7=Extremely negative).

- Baseline vs. TA phase: baseline mean (1.65) > TA mean (1.50). $r(\text{Lag } 1) = .52$.
 - No significant level change: $r = -.15$; $p = .45$.
 - **Significant slope change: $r = .55$; $p = .003$.**
- Baseline vs. FU phase: baseline mean (1.65) > FU mean (1.11). $r(\text{Lag } 1) = .57$.
 - **Significant level change: $r = -.60$; $p = .004$.**
 - No significant slope change: $r = .14$; $p = .57$.

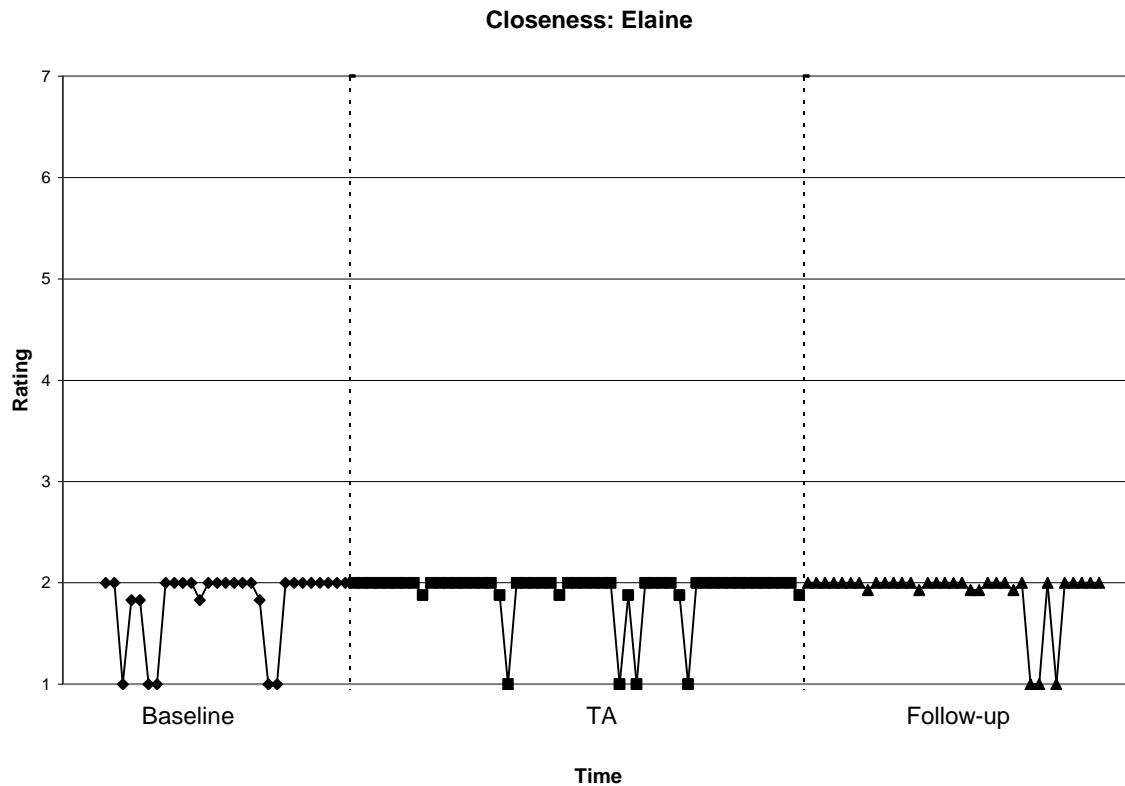
Figure 4.59.



2. How close did you feel to Tom today? (1=Extremely close, 7=Extremely distant).

- Baseline vs. TA phase: baseline (1.80) < TA mean (1.91). $r(\text{Lag } 1) = .22$.
 - No significant level change: $r = .17$; $p = .25$.
 - No significant slope change: $r = .04$; $p = .78$.
- Baseline vs. FU phase: baseline mean (1.80) < FU mean (1.90). $r(\text{Lag } 1) = .33$.
 - No significant level change: $r = .15$; $p = .41$.
 - No significant slope change: $r = .24$; $p = .17$.

Figure 4.60.



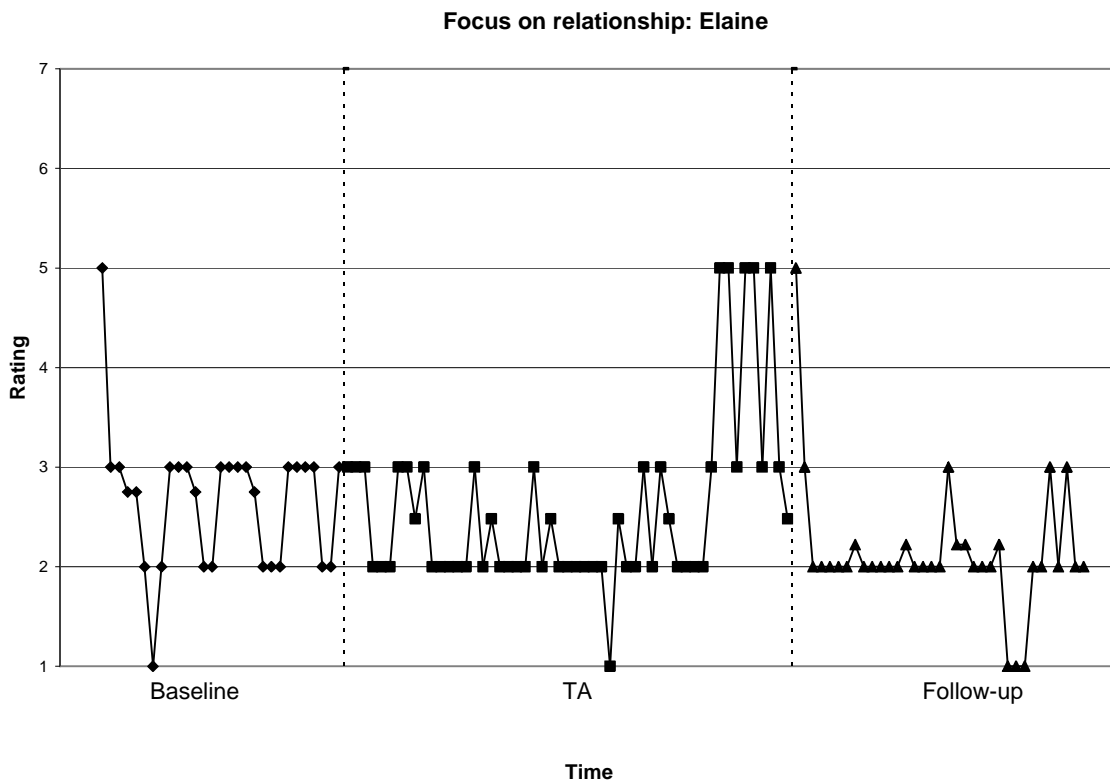
3. How hopeful did you feel about your relationship today? (1=very hopeful, 5=not at all hopeful)

- Time-series analyses not applicable, as she responded “1” (very hopeful) every day throughout the study.

4. To what degree did you feel your attention was focused on your relationship with Tom, versus other things happening in your life (e.g., things going on at work or with other family members or in-laws)? (1= Focused almost exclusively on our relationship; 7= Almost exclusively focused on other things).

- Baseline v. TA phase: baseline mean (2.66) > TA mean (2.58). $r(\text{Lag } 1) = .49$.
 - No significant level change: $r = -.04$; $p = .83$.
 - No significant slope change: $r = -.29$; $p = .12$.
- Baseline vs. FU phase: baseline mean (2.66) > FU mean (2.15). $r(\text{Lag } 1) = .45$.
 - **Nearly significant level change: $r = -.35$; $p = .07$.**
 - No significant slope change: $r = .13$; $p = .53$.

Figure 4.61.

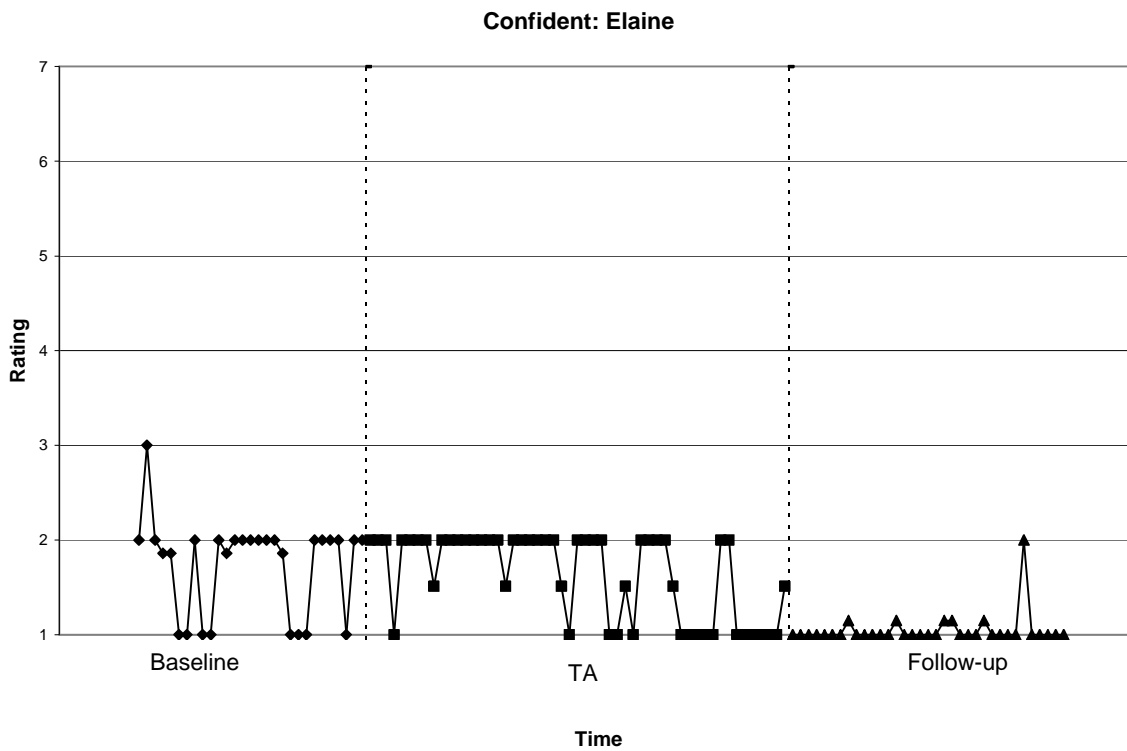


5. On a scale of 1-7, to what degree did you feel confident that you have a strong relationship, or that your relationship is based on a solid foundation? (1=extremely confident; 7=not at all confident).

- Baseline vs. TA phase: baseline mean (1.74) > TA mean (1.64). $r(\text{Lag } 1) = .47$.
 - No significant level change: $r = -.10$; $p = .59$.
 - **Significant slope change: $r = .40$; $p = .03$.**

- Baseline vs. FU phase: baseline mean (1.74) > FU mean (1.05). $r(\text{Lag } 1) = .59$.
 - **Significant level change: $r = -.69$; $p < .001$.**
 - No significant slope change: $r = .03$; $p = .91$.

Figure 4.62.



Summary of time-series results: Elaine

Elaine's ratings of her general feelings toward Tom, as well as her confidence in the relationship, significantly changed in slope during the TA phase, with her ratings showing a gradual increase in confidence and positive feelings over the course of the phase. In addition, the level of her ratings on these two variables improved significantly when comparing the baseline and follow-up phases. Her anxiety ratings also improved significantly when comparing the baseline to the follow-up phase. She also reported feeling more focused on her relationship with Tom during the follow-up phase when compared to the baseline phase, and this difference approached statistical significance.

Clients' perceptions of couples therapy

Visual analysis was used to determine whether the TA was associated with improvements in Tom and Elaine's ratings of their weekly sessions with their couples therapist, Carla. The means and standard deviations of their Session Evaluation Questionnaire—Client Form (SEQ-C) scores are presented in Table A9 (Appendix K), and the graphs of their weekly SEQ-C ratings are presented in Figures 4.64 through 4.71. Again, the SEQ-C was used to assess each partner's feelings about each weekly session with Carla throughout the three phases of the study. Tom neglected to complete the SEQ-C after the last therapy session in the follow-up phase; otherwise, both partners consistently completed the measure after each session with Carla. Tom and Elaine consistently rated sessions as more "deep" at the end of the TA phase and in the first two sessions of the follow-up phase. Tom's smoothness and positivity ratings were also lower in the follow-up phase than in the other two phases.

Figure 4.64.

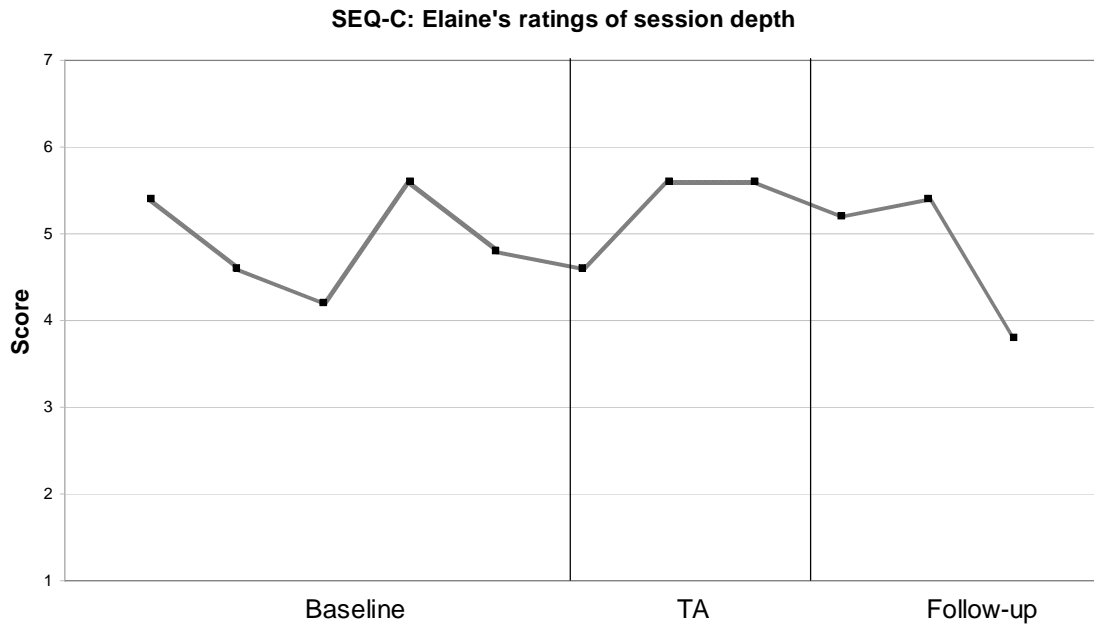


Figure 4.65.

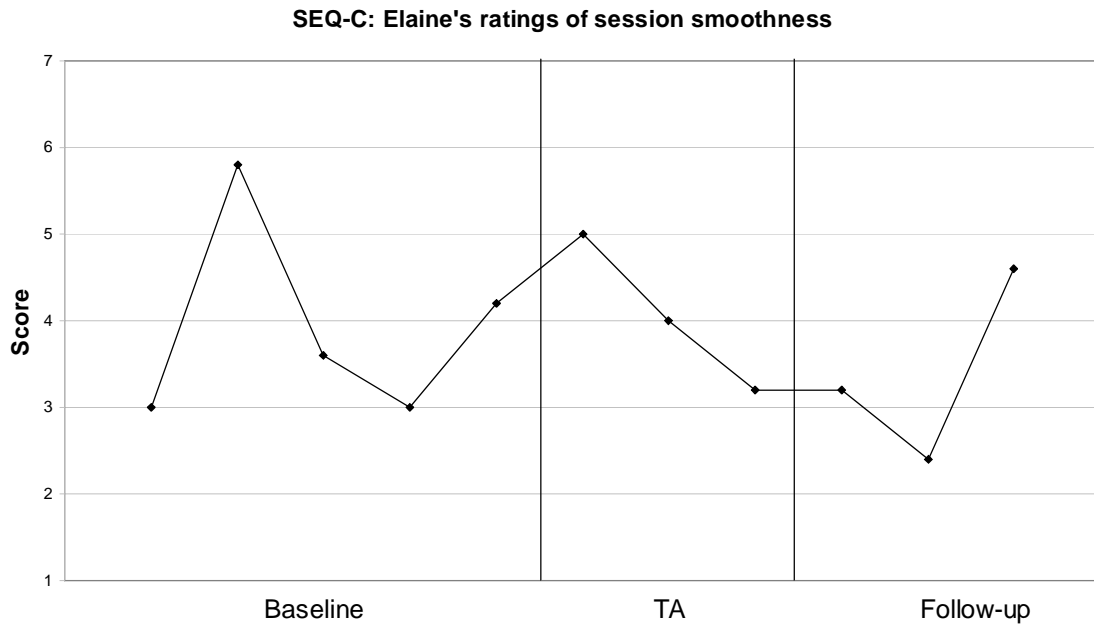


Figure 4.66.

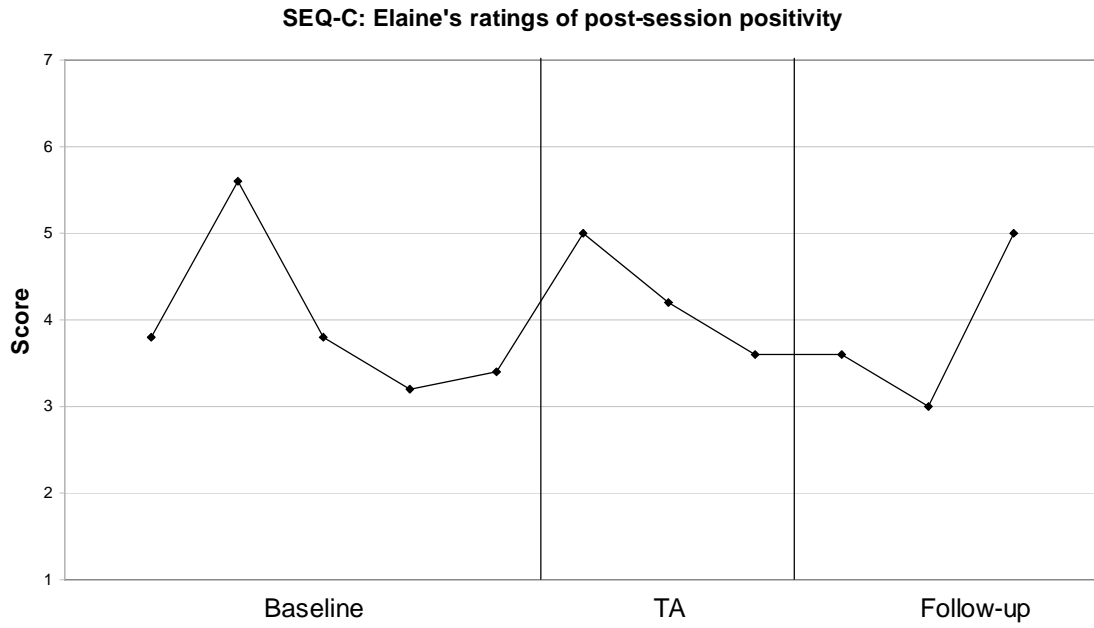


Figure 4.67.

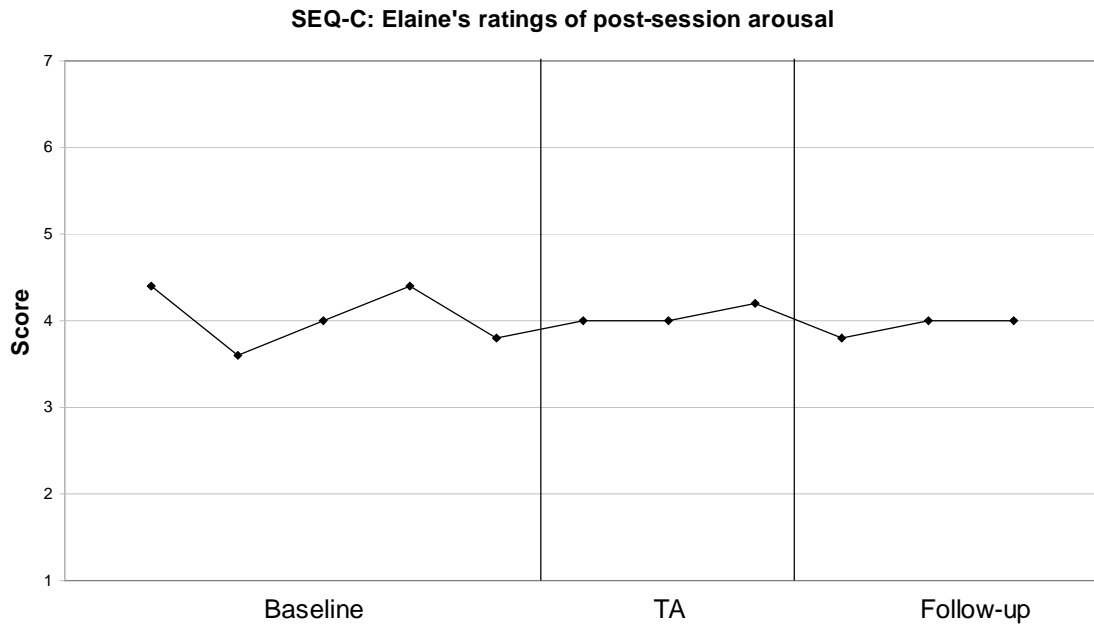


Figure 4.68.

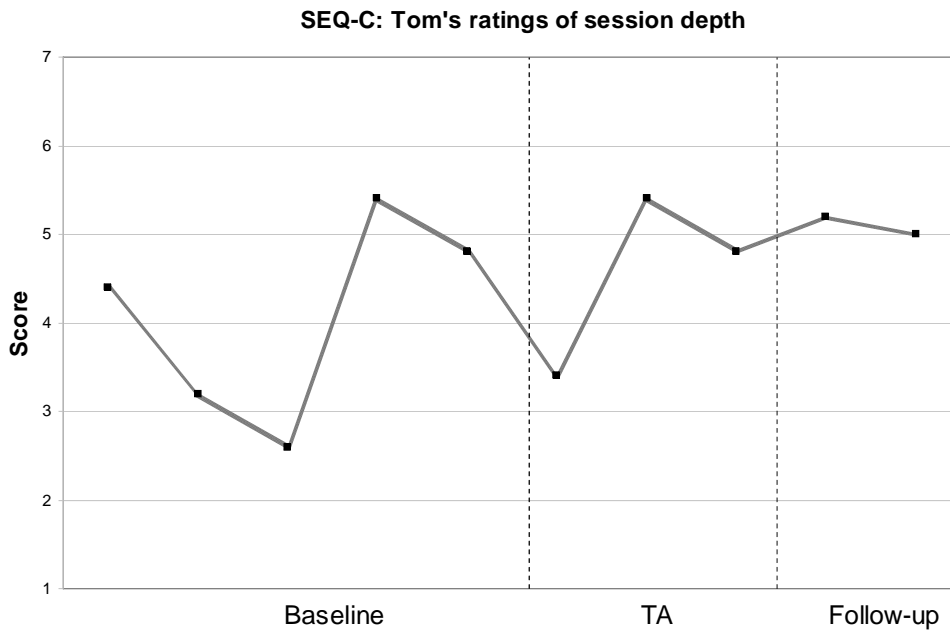


Figure 4.69.

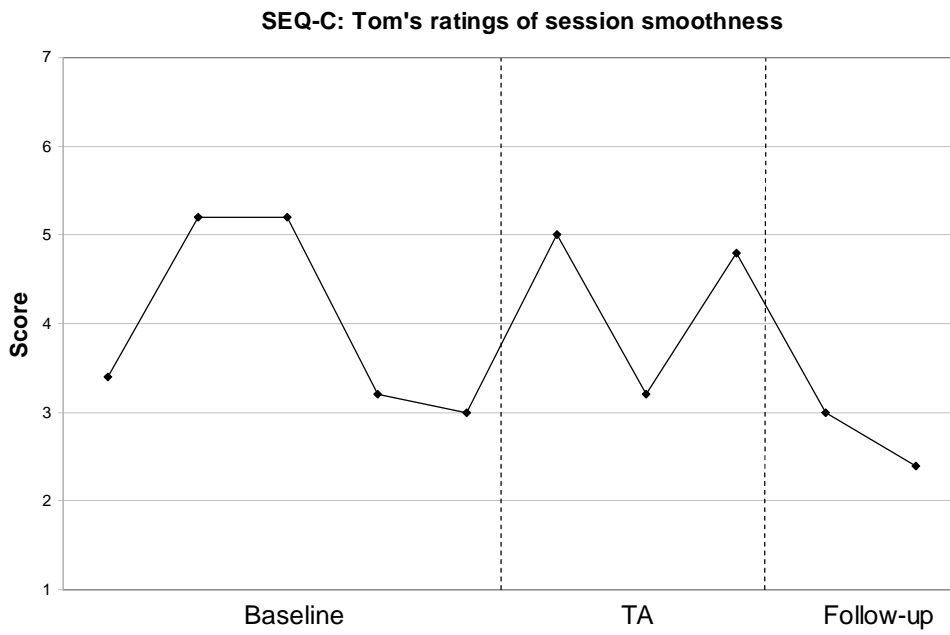


Figure 4.70.

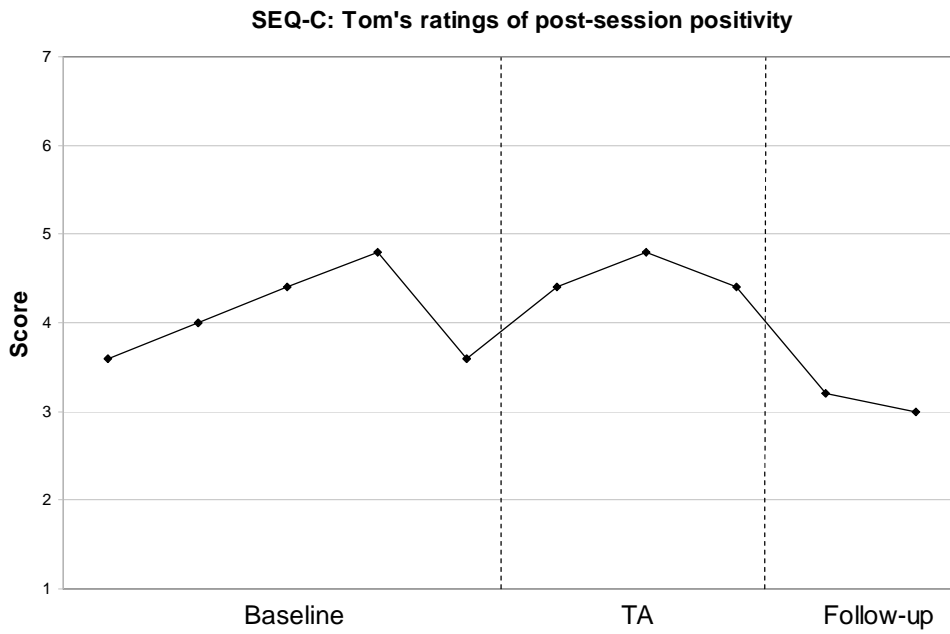
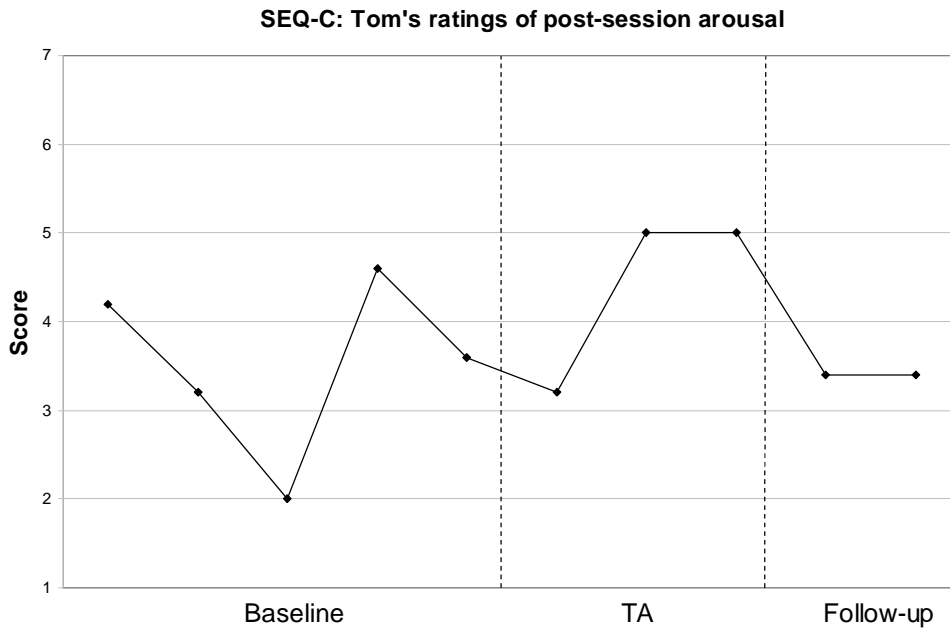


Figure 4.71.



Marital satisfaction and psychological symptoms

The couple's BSI, MDS, and DAS ratings are presented in Table A12, and in Figures 4.72, 4.73, and 4.74. Elaine's Brief Symptom Inventory (BSI) scores decreased by a full standard deviation from Time 1 to Time 2. Her Time 3 score was half a standard deviation below her score at Time 1, and all three of her BSI T-scores fell in the non-distressed range of symptomatology. Tom's BSI scores increased slightly (by less than half a standard deviation) from Time 1 to Time 2. His Time 3 score was nearly a standard deviation lower than his Time 1 score. All three of his BSI T-scores fell within the clinically non-distressed range.

Elaine's marital satisfaction, as measured by the Dyadic Adjustment Scale (DAS), remained more or less constant over the course of the study, fluctuating by no more than four points across phases. According to the cutoff scores that are generally used for the DAS (Sabourin, Valois, & Lussier, 2005), all three of her scores fell in the "absent of significant distress" range. Tom's DAS scores improved somewhat across the course of the study, and all three of his scores fell in the "absent of significant distress" range.

Similar to her DAS scores, Elaine's scores on the Marital Disaffection Scale (MDS) fluctuated very little over the course of the study, and all three fell in the "low" range of disaffection. Tom's MDS scores moved from the "below average" range at Time 1 to the "low" range at both Time 2 and Time 3.

Figure 4.72.

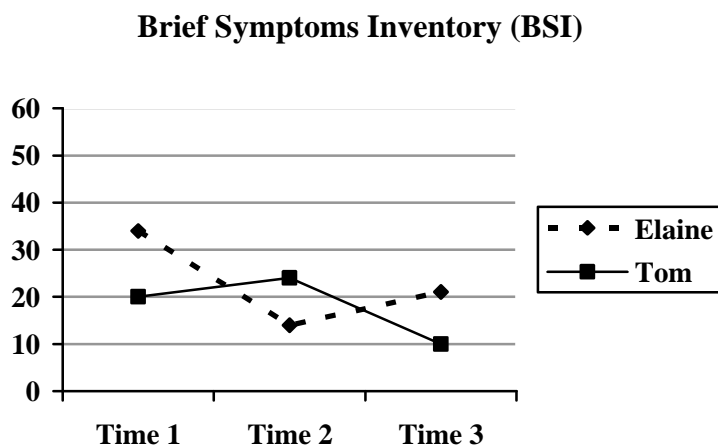


Figure 4.73.

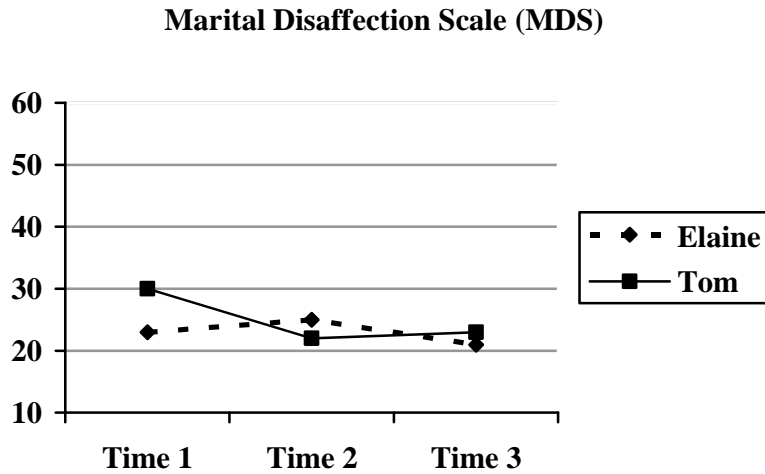
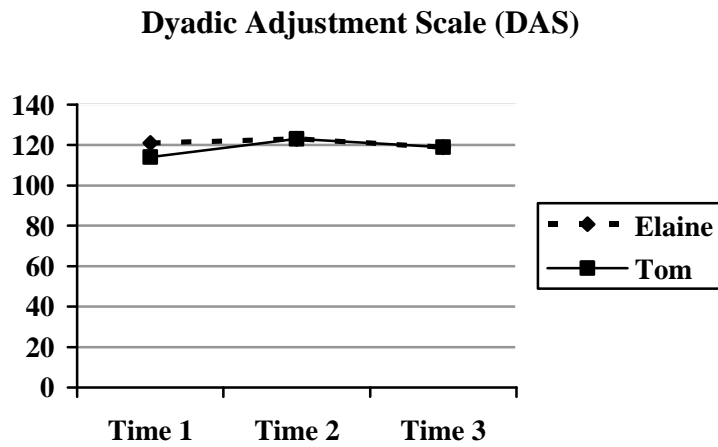


Figure 4.74.



Couple's perceptions of Therapeutic Assessment

Tom and Elaine's T-scores on each of the subscales of the AQ-2 are presented in Table 4.6. On each of the subscales, both partners' scores fell within half a standard deviation of the mean, indicating their satisfaction with the TA was average when compared to most clients receiving psychological assessments at the Center for Therapeutic Assessment.

Table 4.6. *AQ-2 Results (Mean ratings and T-scores)*

	New self-awareness/self-understanding	Positive accurate mirroring	Positive relationship with the examiner	Negative feelings about the assessment	Total satisfaction
Elaine	Mean=3.92 T=51.1	Mean=3.75 T=51.7	Mean=4.17 T=53	Mean=2.42 T=54	Mean=3.81 T=51.9
Tom	Mean=3.92 T=51.1	Mean=3.67 T=51.1	Mean=3.67 T=47.4	Mean=2.33 T=53.5	Mean=3.69 T=50.8

Note: Participants rated each item on a 1-5 Likert-type scale, with high scores indicating higher satisfaction, with the exception of the “negative feelings about the assessment” scale, on which higher ratings indicated more negative feelings.

Couple’s qualitative feedback about the TA

Elaine:

On the whole, Elaine’s qualitative feedback about the Therapeutic Assessment was positive. She reported the assessment met her expectations “fairly well” and commented that it was “a good experience.” She reported that “the MMPI results and the ink blot exercise were most valuable to helping Tom and I understand each other and assess how we communicate with one another.” When asked about the “least valuable” parts of the assessment, she replied, “I didn’t enjoy the memories of spouse - I’m sure they were valuable to help the assessor understand us, but I didn’t feel that they were terribly valuable to me/us.” When asked about how the assessment affected her relationship with Tom, she replied, “I think the assessment has given us some new tools to use when we’re going through difficult times - to stop and try to understand the other person’s point of view, and ask about it.” When asked about how the assessment affected her experience in couples therapy, she replied, “We have not had a session since the assessment, but the process throughout has been enlightening.” She said she did not have any suggestions for improving the way the Center for Therapeutic Assessment conducted assessments. She said she would tell a friend considering an assessment “to be ready for a lot of emotional work, but that it’s worth it.”

Tom:

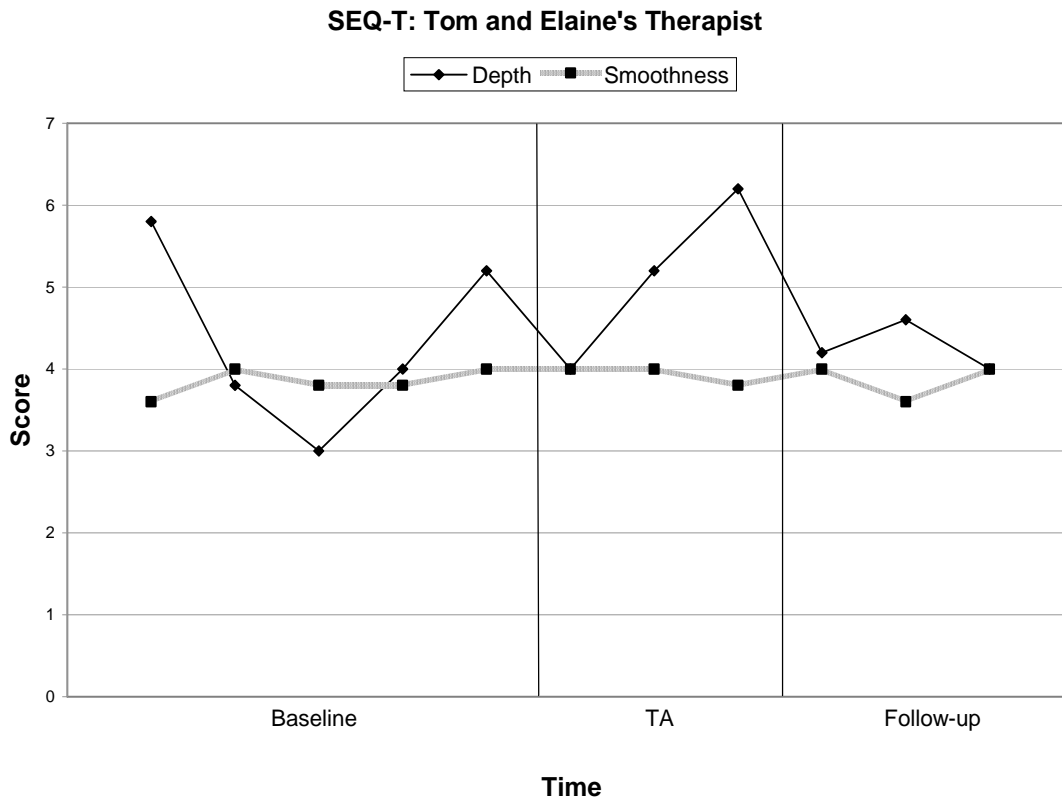
Tom's qualitative responses on the feedback questionnaire were mixed, but largely positive. When asked how well the assessment met his expectations, he replied, "I'm not sure I had a lot of expectations regarding the assessment. But I thought it was a valuable experience." When asked about the "most valuable" parts of the assessment, he replied, "The test where I rated myself and Elaine separately [the FAM-III] was very interesting. I think that showed where we really come together." He reported the memories of spouse exercise was the least valuable part of the assessment, and commented that it "just seemed like a waste of time." When asked how the assessment affected his relationship with Elaine, he replied, "I think it just makes us more aware, especially as it concerns our communication patterns." Regarding how the assessment affected his experience in couples therapy, he stated, "It's hard to say just yet, but I think it will benefit us in relation to some of our personality and communication traits. I think it will make us more aware of these things and allow us to examine our interactions at a deeper level." When asked about suggestions for improving the way assessments are done, he replied, "I think the process took too long. I think it would be more valuable if you had a much more dedicated time period to collect all of the data." He said he would tell a friend considering an assessment, "I think it's a valuable exercise. Though I think I can only really speak from the perspective of the couples assessment. It seems an individual assessment would be a different experience."

Therapist's ratings of couples therapy—Tom and Elaine

To assess for changes in Tom and Elaine's therapist's perceptions of couples therapy before, during, and after the TA, her ratings on the Session Evaluation Questionnaire—Therapist Form (SEQ-T) were examined. The means and standard deviations of her SEQ-T scores are presented in Table A15 (Appendix M) and a chart of her weekly SEQ-T ratings is presented in Figure 4.75. The therapist's ratings of the depth of the therapy sessions appeared to fluctuate from week to week, with the "deepest"

session occurring during the TA phase. Her smoothness ratings, by contrast, fluctuated very little.

Figure 4.75.



Therapist's qualitative feedback—Tom and Elaine

On the whole, Elaine and Tom's therapist appeared to have favorable impressions of the Therapeutic Assessment and its effect on her therapy with the couple. She reported the following as her "hopes and expectations" when she referred the couple for a psychological assessment:

That each member of the couple would gain a greater understanding of themselves, be able to use the information to dive deeper into their emotions as well as identify parts of their emotions and lack of affects that hold them emotionally back.

She stated that the assessment did meet her expectations. She reported the feedback session was the most useful part of the assessment for her, and noted that “it helped the clients identify patterns they enter into and how those patterns might limit their ability to fully express themselves with in the relationship.” She reported being “unsure” about the “least useful” parts of the assessment. When asked what would have made the assessment more useful, she replied, “The clients got a summary of the assessment and I did not. It would be helpful to have the written information.” It should be noted, however, that the assessor had sent a written summary to the therapist, but the report was lost in the mail. The therapist was eventually sent another copy of the report. As a final comment, Tom and Elaine’s therapist stated, “I would be happy to refer clients or colleagues for assessment. I appreciate that the process is as important as the actual tests used in the assessment.”

Discussion: Tom and Elaine

Both Tom and Elaine demonstrated significant improvement on several of their daily ratings from baseline to follow-up. In some cases, such as Tom’s ratings of his focus on the relationship and Elaine’s ratings of her own anxiety, these improvements very well may be clinically meaningful. In other cases, however, it is unclear whether statistically significant changes in ratings represented clinically significant changes in the couple’s lives. From the beginning of the study, some of their daily ratings suffered from ceiling effects, in which there was little room for improvement. For example, when rating her confidence in the relationship, Elaine’s ratings moved from “very confident” (on average) during the baseline phase to “extremely confident” (on average) during the follow-up phase. Similarly, there was little room for improvement on both partners’ scores on the standardized measures of marital satisfaction, as both partners’ scores on the MDS and DAS fell in non-distressed ranges at all three time points. This was also the case with both partners’ BSI scores, which did decrease from baseline to follow-up, but were never in the clinically distressed range.

Tom's daily ratings of his own irritability actually increased significantly during the TA phase, but returned to baseline levels during the follow-up phase. This is similar to the pattern seen in some of Susan and George's daily ratings, and again indicates that although the TA may have "stirred up" unpleasant feelings in Tom, he was able to recover from, and perhaps even benefit from this temporary emotional disruption.

As with the other two couples, it is difficult to determine what sort of effect, if any, the TA had on the couple's experience in therapy with Carla, their couples therapist. For Tom, Elaine, and Carla, some of the "deepest" sessions occurred during the TA phase, and Tom and Elaine also rated first two sessions in the follow-up phase as being relatively deep. Again, however, it is impossible to know whether this represented a lasting deepening of therapy sessions, a temporary reaction to the TA, or something else altogether.

In many ways, Tom and Elaine were quite different from the other two couples in that there was very little overt conflict between the partners, and each partner consistently reported having positive feelings toward the other one. Rather than changing the couple's feelings about one another, the TA appears to have affected individual-level factors (e.g., Elaine's anxiety) and external factors (e.g., each partner's tendency to get wrapped up in stressors related to their families of origin). It is important to remember that, compared to the other two couples, Tom and Elaine had only been together for a short time, and did not suffer from longstanding marital conflicts. Given these factors, it is possible that, through helping the couple become more focused on one another and, in Elaine's case, less anxious, the TA may have had a preventative effect, rather than a curative effect. In other words, the intervention may have helped equip the couple with tools that might help them cope with future stressors (e.g., conflict with other family members) while maintaining a healthy focus on their relationship with one another. However, this is only speculative, and longer-term follow-up data would be necessary in order to assess the long-term health of the marriage.

Chapter V: Summary and Conclusions

Summary of results from the three cases

Although the results of the current study were described in three individual case studies, each with varying results, some generalizations can be made across cases. In the following section, the study results will be summarized as they relate to each of the six hypotheses proposed at the outset of the study.

Hypothesis 1: A. After participating in the TA intervention, participants will report improvements on ideographic (individualized), daily measures of relationship satisfaction. B. These improvements in relationship satisfaction will be sustained, or will continue to increase at follow-up.

Speaking at the most general level, each participant demonstrated changes for the better, either in the level or slope of their ratings, on some of the daily ratings of marital satisfaction. Because results on the daily measures varied a great deal from couple to couple, I will summarize the time-series results by couple here.

Couple #1 (Mary and Ray): Recall that there were three standard, daily questions to which all participants were asked to respond. On the first of these questions (“general feelings about your partner”), Ray and demonstrated nearly significant level changes from baseline to follow-up, but Mary did not. On the second standard daily question (“How close did you feel to your partner?”) Mary demonstrated a nearly significant change in slope when comparing her ratings at baseline to those at both the TA and follow-up phases. Ray demonstrated a significant change in slope during the TA phase, but this change was not sustained at follow-up. On the third standard question (“hopefulness about the relationship), Mary’s ratings changed significantly in slope from baseline to follow-up. Although Ray’s ratings demonstrated a significant change in slope during the TA phase, these changes were not sustained at follow-up. Results for the

individualized daily measures also varied. Ray's ratings improved significantly or nearly significantly (in either level or slope) from baseline to follow-up on three of his four individualized measures. The slope of Mary's ratings made significant or nearly significant changes for the better from baseline to follow-up on four of the five individualized measures.

Couple #2 (George and Susan): On the first of the three standard daily questions ("general feelings about your partner"), both George and Susan demonstrated significant or nearly significant changes (in either level or slope) during the TA phase, but these changes were not sustained at follow-up. On the second standard question ("How close did you feel to your partner?"), both partners demonstrated significant or nearly significant changes in level during the TA phase, but these changes were not sustained at follow-up. On the third standard question ("hopefulness about the relationship), Susan's ratings changed significantly in slope from baseline to follow-up. Although George demonstrated changes in both slope and level during the TA phase, these changes were not sustained at follow-up. For both George and Susan, their ratings improved significantly (in level) from baseline to follow-up on only one of the four individualized measures. In addition, both partners actually reported changes for the worse on one measure during the TA phase; however, in both cases, the changes returned to (but did not improve upon) baseline levels during the follow-up phase.

Couple #3 (Elaine and Tom): On the first of the three standard daily questions ("general feelings about your partner"), Elaine demonstrated significant level changes from baseline to follow-up, but Tom's ratings made no significant changes across the course of the study. On the second standard question ("How close did you feel to your partner?") Tom demonstrated a significant level change from baseline to follow-up. Elaine did not show any significant changes on this variable across the course of the study. Similarly, on the third standard question

(“hopefulness about the relationship), Tom’s ratings demonstrated a significant level change from the baseline to the TA phase, and this change was still nearly significant at follow-up. Again, Elaine’s ratings on this measure did not change significantly during the study. Tom’s ratings improved significantly (in level) from baseline to follow-up on only one of the three individualized measures. Elaine’s ratings made significant or nearly significant changes in level from baseline to follow-up on all three of her three daily measures. In addition, Tom reported changes for the worse on one measure during the TA phase; however, his ratings returned to (but did not improve upon) baseline levels during the follow-up phase.

Hypothesis 2: A. After participating in the TA intervention, participants will report increases in relationship satisfaction as measured by nomothetic (global) measures of relationship functioning. B. These increases will be sustained or will continue to improve at follow-up.

On the whole, most participants reported improved relationship satisfaction on both the Marital Disaffection Scale (MDS) and the Dyadic Adjustment Scale (DAS) over time. Four of the six participants demonstrated improved MDS scores (with improvement sustained at follow-up), while the other two participants’ scores remained constant across phases. For the two participants who showed no improvement on the MDS, their scores at the outset of the study were already low, indicating low baseline levels of “disaffection” with very little room for improvement.

Similarly, four of the six participants demonstrated improved DAS scores, again with improvement sustained at follow-up. Again, the two participants who showed no improvement on the DAS had reported high levels of marital satisfaction on this measure at baseline, leaving little room for improvement in their scores.

Hypothesis 3: A. Participants will report improved ratings of progress in couples therapy. B. Couples' therapists will also report improved ratings of their clients' progress in therapy.

For the most part, judging from the Session Evaluation Questionnaire (SEQ) scores, it was difficult to discern whether there were significant changes in both the clients' and therapists' perceptions of couples therapy over time. In some cases, the clients and their therapists rated sessions occurring during the TA and/or follow-up phases as relatively "deep" compared to sessions in the baseline phase; however, these differences in ratings were not dramatic, and it is impossible to know whether any deepening that occurred was sustained beyond the four-week follow-up period.

Hypothesis 4: A. Participants will report decreases in psychological symptoms after participating in the TA intervention. B. These decreases in symptomatology will be sustained, or will continue to decrease at follow-up.

In two cases, the women reported decreased symptomatology on the Brief Symptom Inventory (BSI), while the men reported slightly increased symptomatology from baseline to follow-up. In the third case, both partners reported decreased symptomatology from baseline-to follow-up.

Hypothesis 5: Participants will report feeling satisfied with the TA intervention. Specifically, participants will report above average levels of satisfaction on all four subscales of the AQ-2 (T-scores above 50). In addition, it is hypothesized that, overall, participants' written feedback on the Feedback Questionnaire will be positive.

On the whole, participants reported average levels of satisfaction with the TA on the Assessment Questionnaire (AQ-2). Given that the AQ-2 was normed on clients at the Center for Therapeutic Assessment, the scores in the current sample fell in the expected range. The highest T-score (57) was on the "positive relationship with the examiner" scale, but was still within one standard deviation of the mean. Qualitative feedback from participants was largely positive, and most participants reported the TA provided them

with greater insight into themselves and their relationship. For the most part, negative feedback focused on specific aspects of the TA; for example, one couple stated they did not think the memories of spouse procedure was worth their while. Another participant said he wished he had been given more information upfront about what the TA would entail. Another participant, who had been in couples therapy for three years at the time of the study, stated that the answers to her assessment questions were not as valuable as she had hoped they would be, and noted that the assessment might have been more helpful at an earlier stage of couples therapy.

Hypothesis 6: Couples' therapists will report feeling satisfied with the TA intervention. Specifically, therapists will report feeling that the intervention aided the couples' therapy.

All three therapists provided positive feedback about the assessment, and seemed to believe the TA had been useful to them in their work with the couples. All three said they would recommend the intervention to colleagues. One therapist's feedback indicated she would have preferred a more comprehensive TA for the couple, with individual Rorschachs for each client, more information on the clients' MMPI-2 scores, and multiple feedback sessions.

Discussion

The participants in the current study reported some positive significant changes in their relationship satisfaction and symptomatology, although there were no sweeping, dramatic results across the board. In retrospect, when looking at the context of the study, it becomes clear that the three TA assessors were given a particularly difficult test to pass. First, the study used an abbreviated TA model in which each assessor administered a predetermined set of assessment instruments. This prevented the assessors from choosing assessment tools based on the clients' questions and presenting problems (an important aspect of TA), and also limited the amount of information that the assessors could gather from each client. In addition, all three couples had been in therapy for a

while (range: 9 months to 3 years) at the outset of the study, and all three therapists (particularly George and Susan's therapist) were well-regarded in the community as expert couples therapists. Based on initial interviews with couples, it appears that all three couples had already made important gains in couples therapy, and it follows that many of the changes they were now hoping to make would likely be more subtle in nature.

Despite all of the above-mentioned complicating factors, it does appear that the TA may have helped some of the participants with their "stuckness." Although one can only speculate, it appears this may have been the case with Ray and Mary, who were able to disclose and process an infidelity during the TA. In other cases, the progress made as a result of the TA is less clear. For example, the TA was not successful in helping George and Susan resolve problems in their sex life (at least not by the end of the follow-up phase), and it may be that the brief intervention was simply not powerful enough to change such a longstanding problem in such a short period of time. In addition, in some cases (particularly with Tom and Elaine, who were experiencing very low levels of dissatisfaction at the outset of the study), the statistically significant changes on daily measures may not have reflected clinically significant changes (e.g., going from feeling "very hopeful" about the relationship to "extremely hopeful").

For a few of the participants, the TA appeared to be associated with substantial decreases in symptomatology. This is an interesting finding, since improved individual-level functioning was thought to be only a secondary gain in couples TA. This finding also makes intuitive sense in a way, however, as two of the main assessment tools that were used (the Rorschach and the MMPI-2) assess individual-level functioning, and information from these assessments would likely help clients make improvements at an individual level, in addition to the couple level. The fact that two of the men developed *more* symptoms over the course of the study, while their wives improved, is another interesting finding. As was mentioned in the previous chapter, it seems possible that TA may have altered the equilibrium in each of these relationships so that one partner was no longer "holding" all of the pathology in the couple. It may also be the case that the TA

“stirred up” painful emotions for these men, thus leading to increases in reported symptoms. Unfortunately, due to the short-term follow-up period, we cannot know the degree to which the changes in symptomatology were temporary versus lasting. In addition, further research is necessary to better understand the mechanisms of symptom change in a couples assessment.

Qualitatively, all three couples and their therapists gave largely positive feedback about the TA and believed it fostered new and useful insights. Much of the negative qualitative feedback seemed to be partly a result of having used a standard protocol for the purposes of the research study (as opposed to an individually-tailored TA protocol). For example, each participant and couple seemed to identify different aspects of the TA as being the “most” or “least” helpful. In a more typical TA, assessors would have chosen instruments that they believed to be a good fit for the clients’ presenting questions. In addition, one therapist said she wished the TA had been more comprehensive, while one participant said he thought the process took too long. It would be interesting, therefore, to know whether TA interventions that are more specifically tailored to couples’ presenting concerns and questions lead to better outcomes, both qualitative and quantitative.

Strengths, Limitations, and Future Directions

There are many unique strengths and limitations to the time-series design used in the current study. First, the study’s small sample size makes it difficult to generalize results. As we have seen in the above section, results were quite variable across and within couples, and it is difficult to say why the TA appeared to have a positive effect on some variables but not others. Future studies using larger sample sizes could provide valuable information about the effectiveness of TA for the larger population of distressed couples. In addition, because the daily questions were developed specifically for this study, and most questions were tailored specifically to participants’ presenting problems, the reliability and validity of these measures was unknown. Furthermore, the daily questions asked couples to rate their subjective experience (e.g., the degree to which they

got caught up in a “power struggle), rather than clear, observable behaviors, which made valid and reliable measurement difficult.

At the same time, however, the single case design allowed for a much more in-depth look at the day-to-day fluctuations in marital satisfaction than a large-n study would have permitted. For example, large spikes on some of Mary and Ray’s daily ratings can be linked directly to the disclosure of Ray’s affair, and it was also possible to observe the fallout and aftershocks from this event in each partner’s daily ratings. In a large randomized controlled trial, this important information would have been lost.

At the outset of the current study, all three couples had been in therapy for a year or more and had reported making significant gains in couples therapy. As a result, there was likely not as much “room for improvement” as there might have been if couples had been referred to the TA at the height of their marital distress. In addition, the fact that the participants were in ongoing couples therapy throughout the study presented a possible confound. Although the use of a four-week baseline phase helped support the conclusion that significant improvement in dependent variables was the result of the TA intervention, it is impossible to know for certain whether these improvements were caused by the TA, the couples therapists’ interventions, or some combination of the two. Future studies examining the utility of TA as an intervention for couples who are not in therapy, or who are at an earlier stage of couples therapy, will be necessary.

The use of a one-month follow-up period may also represent a limitation, as it only provided information about clients’ short-term gains in marital satisfaction and symptomatology. A longer follow-up period would have provided valuable information about the durability of therapeutic gains. It also would have allowed the researcher to assess whether certain variables that had not improved during the one-month follow-up period began to improve later. Indeed, Finn (2007) has noted that many clients do not begin to fully assimilate assessment findings until a full month (or more) after the conclusion of the TA. A longer follow-up period also would have allowed the researcher to observe any potential clinical gains that were made once the couples’ therapists had had more time to incorporate the TA findings into their clinical formulations and

treatment plans. An extended follow-up period would therefore be an important component of future studies of TA with couples.

Appendix A: Demographic Questionnaire

Birth date (dd/mm/yy): _____

Sex (CIRCLE ONE): Male Female

Ethnicity/Race (CIRCLE ONE):

European American/White	African American/Black
Asian American/Pacific Islander	Indian American
Native American	Latino/Latina
Multiracial	Other: _____
Prefer not to answer	

Highest level of education (CIRCLE ONE):

High school degree/GED
Some college
Professional degree
Bachelor's degree
Graduate degree

What is your current profession? _____

How long have you and your partner been married/living together? _____

Do you have children? (circle one) Yes No

If you do have children:

How many children do you have with your current partner? _____

How many children do you have from a previous relationship _____

How many of these children currently live with you? _____

Appendix B: Dyadic Adjustment Scale (DAS)

Most persons have disagreements in their relationships. Please indicate below the approximate extent of agreement or disagreement between you and your partner for each on the following list.

	Always Agree	Almost Always Agree	Occasionally Agree	Frequently Disagree	Almost Always Disagree	Always Disagree
1. Handling family matters	5	4	3	2	1	0
2. Matters of recreation	5	4	3	2	1	0
3. Religious matters	5	4	3	2	1	0
4. Demonstrations of affection	5	4	3	2	1	0
5. Friends	5	4	3	2	1	0
6. Sex relations	5	4	3	2	1	0
7. Conventionality (correct or proper behavior)	5	4	3	2	1	0
8. Philosophy of life	5	4	3	2	1	0
9. Ways of dealing with parents or in-laws	5	4	3	2	1	0
10. Aims, goals, and things believed important	5	4	3	2	1	0
11. Amount of time spent together	5	4	3	2	1	0
12. Making major decisions	5	4	3	2	1	0
13. Household tasks	5	4	3	2	1	0
14. Leisure time interests and activities	5	4	3	2	1	0
15. Career decisions	5	4	3	2	1	0
	All the time	Most of the time	More often than not	Occasionally	Rarely	Never
16. How often do you discuss or have you considered divorce, separation, or terminating your relationship?	0	1	2	3	4	5
17. How often do you or your mate leave	0	1	2	3	4	5

the house after a fight?						
18. In general, how often do you think that things between you and your partner are going well?	5	4	3	2	1	0
19. Do you confide in your mate?	5	4	3	2	1	0
20. Do you ever regret that you married (or lived together)?	0	1	2	3	4	5
21. How often do you and your partner quarrel?	0	1	2	3	4	5
22. How often do you and your mate get on each other's nerves?	0	1	2	3	4	5
	Every day	Almost every day	Occasionally	Rarely	Never	
23. Do you kiss your mate?	4	3	2	1	0	
	All of them	Most of them	Some of them	Very few of them	None of them	
24. Do you and your mate engage in outside interests together?	4	3	2	1	0	
How often would you say the following events occur between you and your mate?						
	Never	Less than once a month	Once or twice month	Once or twice a week	Once a day	More often
25. Having a stimulating exchange of ideas.	0	1	2	3	4	5
26. Laugh together	0	1	2	3	4	5
27. Calmly discuss something	0	1	2	3	4	5
28. Work together on a project	0	1	2	3	4	5
These are some things about which couples sometimes agree and sometimes disagree. Indicate if either item below caused differences of opinions or were problems in your relationship during the past few weeks. (Check yes or no)						
	Yes		No			
29. Being too tired for sex	0		1			
30. Not showing love	0		1			

31. The dots on the following line represent different degrees of happiness in your relationship. The middle point, “happy,” represents the degree of happiness in most relationships. Please circle the dot which best describes the degree of happiness, all things considered, of your relationship.

0	1	2	3	4	5	6
Extremely <u>Un</u> happy	Fairly <u>Un</u> happy	A little <u>Un</u> happy	Happy	Very happy	Extremely happy	Perfect

32. Which of the following statements best describes how you feel about the future of your relationship?

 1 I want desperately for my relationship to succeed, *and would go to almost any length* to see that it does.

 2 I want very much for my relationship to succeed, and *will do all I can* to see that it does.

 3 I want very much for my relationship to succeed, and *will do my fair share* to see that it does.

 4 It would be nice if my relationship succeeded, *I can't do much more than I am doing now* to help it.

 5 It would be nice if it succeeded, but *I refuse to do any more than I am doing now* to keep the relationship going.

 6 My relationship can never succeed, and *there is not more that I can do* to keep the relationship going.

Appendix C: Marital Disaffection Scale (MDS)

Key: Very True=4
Somewhat True=3
Not Very True=2
Not at All True=1

1. If I could never be with my spouse, I would feel miserable.
2. I find it difficult to confide in my spouse about a number of things.
3. I enjoy spending time alone with my spouse.
4. I often feel lonely even though I am with my spouse.
5. I miss my spouse when we're not together for a couple of days.
6. Most of the time I feel very close to my spouse.
7. I seem to enjoy just being with my spouse.
8. I look forward to seeing my spouse at the end of the day.
9. My love for my spouse has increased more and more over time.
10. I find myself withdrawing more and more from my spouse.
11. When I have a personal problem, my spouse is the first person I turn to.
12. Apathy and indifference best describe my feelings toward my spouse.
13. I feel little, if any, desire to have sex with my spouse.
14. My spouse has always been there when I needed him or her.
15. I would prefer to spend less time with my spouse.
16. I have more positive than negative thoughts about my partner.
17. I have a lot of angry feelings toward my spouse.
18. I am not as concerned about fulfilling my obligations and responsibilities in my marriage as I was in the past.
19. I try to avoid spending time with my spouse.
20. There are times when I do not feel a great deal of love and affection for my mate.
21. I enjoy sharing my feelings with my spouse.

Appendix D: Brief Symptom Inventory (BSI)

Instructions: Below is a list of problems and complaints that people sometimes have. For each one, tell me how much that problem has bothered or distressed you during the past week, including today. Please tell me whether each problem has bothered you not at all, a little bit, moderately, quite a bit, or extremely.

	Not at all	A little bit	Moderately	Quite a bit	Extremely
1. Nervousness or shakiness inside.					
2. Faintness or dizziness.					
3. The idea that someone else can control your thoughts.					
4. Feeling others are to blame for most of your troubles.					
5. Trouble remembering things.					
6. Feeling easily annoyed or irritated.					
7. Pains in heart or chest.					
8. Feeling afraid in open spaces.					
9. Thoughts of ending your life.					
10. Feeling that most people cannot be trusted.					
11. Poor appetite.					
12. Suddenly scared for no reason.					
13. Temper outbursts that you could not control.					
14. Feeling lonely even when you are with people.					
15. Feeling blocked in getting things done.					
16. Feeling lonely.					
17. Feeling blue.					
18. Feeling no interest in things.					
19. Feeling fearful.					
20. Your feelings being easily hurt.					
21. Feeling that people are unfriendly or dislike you.					
22. Feeling inferior to others.					
23. Nausea or upset stomach.					
24. Feeling that you are watched or talked about by others.					
25. Trouble falling asleep.					
26. Having to check and double check what you do.					
27. Difficulty in making decisions.					

28. Feeling afraid to travel on buses, subways, or trains.					
29. Trouble getting your breath.					
30. Hot or cold spells.					
31. Having to avoid certain things, places, or activities because they frighten you.					
32. Your mind going blank.					
33. Numbness or tingling in parts of your body.					
34. The idea that you should be punished for your sins.					
35. Feeling hopeless about the future.					
36. Trouble concentrating.					
37. Feeling weak in parts of your body.					
38. Feeling tense or keyed up.					
39. Thoughts of death or dying.					
40. Having urges to beat, injure, or harm someone.					
41. Having urges to break or smash things.					
42. Feeling very self-conscious with others.					
43. Feeling uneasy in crowds.					
44. Never feeling close to another person.					
45. Spells of terror or panic.					
46. Getting into frequent arguments.					
47. Feeling nervous when you are left alone.					
48. Others not giving you proper credit for your achievements.					
49. Feeling so restless you could not sit still.					
50. Feelings of worthlessness.					
51. Feeling that people will take advantage of you if you let them.					
52. Feelings of guilt.					
53. The idea that something is wrong with your mind.					

Appendix E: Assessment Questionnaire-2 (AQ-2)

Instructions

This questionnaire deals with your thoughts and feelings about your psychological assessment. Please read each statement carefully. Once you decide how much you agree or disagree with a statement, circle the number that best matches how the statement applies to you. Be as honest and as accurate as possible. Please do not skip any item and circle only one number for each statement.

Use the following scale to rate each statement:

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	
	1	2	3	4	5	
				Strongly Disagree		Strongly Agree
1. The assessment did not teach me anything new about myself.	1	2	3	4	5	
2. The assessment made me proud of who I am.				1	2	3
3. The assessor earned my respect.				1	2	3
4. I felt I was under a microscope.				1	2	3
5. The assessor introduced me to new aspects of myself.				1	2	3
6. The assessment made me feel good about myself.				1	2	3
7. It was easy to trust the assessor.				1	2	3
8. The assessment hurt me.				1	2	3
9. I gained a new understanding of myself.				1	2	3
10. The assessment captured the “real” me.				1	2	3
11. The assessor seemed to like me.				1	2	3
12. The assessment was unsettling to me.				1	2	3
13. The assessment confirmed parts of me that I had only suspected.				1	2	3
14. The assessor said nice things about me.				1	2	3
15. I felt very close to the assessor.				1	2	3
16. The assessment was a humiliating and degrading experience.	1	2	3	4	5	
17. The assessment made me think of myself.				1	2	3
18. The assessment made me feel important.				1	2	3
19. The assessor treated me warmly.				1	2	3
20. The assessment was emotionally draining.				1	2	3
21. I am more aware of how I behave with other people.				1	2	3

22. I felt special.	1	2	3	4	5
23. I really connected with the assessor.	1	2	3	4	5
24. At times during the assessment, I felt like I did when I was a child.	1	2	3	4	5
25. The assessment helped me organize my thoughts about myself.	1	2	3	4	5
26. The assessment confirmed how I see myself.	1	2	3	4	5
27. I liked the assessor.	1	2	3	4	5
28. The assessment made me feel that my life is nothing but problems.	1	2	3	4	5
29. I have changed the way I think about my problems.	1	2	3	4	5
30. I feel more sure of who I am.	1	2	3	4	5
31. The assessor was interested in what I had to say.	1	2	3	4	5
32. I felt judged by the assessor.	1	2	3	4	5
33. I am more aware of how I am feeling.	1	2	3	4	5
34. I felt my strengths were recognized.	1	2	3	4	5
35. The assessor treated me as an equal.	1	2	3	4	5
36. The assessor made me feel inadequate.	1	2	3	4	5
37. The assessment will make a difference in my upcoming decisions.	1	2	3	4	5
38. The assessment made me think about where I am headed in my life.	1	2	3	4	5
39. I felt that the assessor respected me.	1	2	3	4	5
40. The assessor insulted me.	1	2	3	4	5
41. I am more aware of why people react to me the way they do.	1	2	3	4	5
42. I know that how I see myself is really true.	1	2	3	4	5
43. The assessor and I worked as a team to learn more about me.	1	2	3	4	5
44. I felt exposed.	1	2	3	4	5
45. I can think of myself as I never had before.	1	2	3	4	5
46. The assessment described thoughts & feelings I have about myself	1	2	3	4	5
47. The assessor was on my side.	1	2	3	4	5
48. The assessment made me rethink the way I already viewed myself.	1	2	3	4	5

Appendix F: Feedback Questionnaire—Client Form

1. How well did the assessment meet your expectations?
2. What part(s) of the assessment did you find most valuable?
3. What part(s) of the assessment were least valuable?
4. How do you think the assessment has affected your relationship with your partner?
5. How do you think the assessment has affected your experience in couples therapy?
6. What suggestions do you have for improving the way we do assessments?
7. What would you tell a friend who was considering getting an assessment?
8. Please give any other comments. Use the back of the form if needed.

Appendix G: Feedback Questionnaire—Therapist Form

1. Briefly, what were your hopes and expectations when you referred this couple for a psychological assessment?
2. Did the assessment meet your expectations?
3. What part(s) of the assessment was most useful to you and your clients?
4. What part was least useful?
5. What would have made the assessment more useful?
6. What would you tell a colleague who was considering referring a client to me/us for this type of assessment?

Appendix H: Session Evaluation Questionnaire

Please circle the appropriate number to show how you feel about this session.

This session was:									
	bad	1	2	3	4	5	6	7	good
	difficult	1	2	3	4	5	6	7	easy
	valuable	1	2	3	4	5	6	7	worthless
	shallow	1	2	3	4	5	6	7	deep
	relaxed	1	2	3	4	5	6	7	tense
	unpleasant	1	2	3	4	5	6	7	pleasant
	full	1	2	3	4	5	6	7	empty
	weak	1	2	3	4	5	6	7	powerful
	special	1	2	3	4	5	6	7	ordinary
	rough	1	2	3	4	5	6	7	smooth
	comfortable	1	2	3	4	5	6	7	uncomfortable
Right now I feel:									
	happy	1	2	3	4	5	6	7	sad
	angry	1	2	3	4	5	6	7	pleased
	moving	1	2	3	4	5	6	7	still
	uncertain	1	2	3	4	5	6	7	definite
	calm	1	2	3	4	5	6	7	excited
	confident	1	2	3	4	5	6	7	afraid
	friendly	1	2	3	4	5	6	7	unfriendly
	slow	1	2	3	4	5	6	7	fast
	energetic	1	2	3	4	5	6	7	peaceful
	quiet	1	2	3	4	5	6	7	aroused

Appendix I: Treatment Integrity Rating Forms

Initial session with couple

Please rate the degree to which you agree or disagree with each of the following statements.

	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strongly Agree
1. The assessor sought assessment questions from both clients.					
2. The assessor asked the clients about their “best guess” answers to at least one of these questions.					
3. The assessor let the clients know that he/she would be asking for their input/collaboration throughout the assessment process.					
4. The assessor talked to the clients about the tests they would be taking.					
5. The assessor demonstrated empathy.					
6. The assessor demonstrated good listening skills.					
7. The assessor encouraged the clients to ask questions about the assessment.					
8. If a client asks a non-systemic question about his/her partner, the assessor helped to rephrase this as a systemic question.					

Consensus Rorschach/Assessment Intervention session

Please rate the degree to which you agree or disagree with each of the following statements.

	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strong ly Agree
1. During the Consensus Rorschach task, the assessor asked the couple to come up with responses that they could both see and agree on.					
2. After administering the Consensus Rorschach, the assessor asked the clients to talk about their experience during the task (e.g., thoughts, feelings, observations).					
3. The assessor attempted to make connections between what happened during the Consensus Rorschach task and what happens in other areas of the couple's life together.					
4. The assessor asked the couple to try out a new approach to the Consensus Rorschach (using a card that was not used during the first administration).					
5. The assessor demonstrated empathy.					
6. The assessor demonstrated good listening skills.					
7. The assessor made systemic interpretations about how each person contributes to the patterns revealed.					

Summary/Discussion session

Please rate the degree to which you agree or disagree with each of the following statements.

	1 Strongly Disagree	2 Disagree	3 Neither agree nor disagree	4 Agree	5 Strong ly Agree
1. The assessor appeared to begin by presenting Level 1 findings, followed by Level 2 and (if appropriate) Level 3 findings.					
2. The assessor encouraged the clients to revise or amend the findings.					
3. The assessor demonstrated empathy.					
4. The assessor demonstrated good listening skills.					
5. The assessor seemed to give equal emphasis to each partner's contribution to the relationship struggles.					
6. The assessor asked clients to give real life examples of assessment findings that were discussed.					

Appendix J: Means, standard deviations, and slopes for daily questions

Table A1. Means, standard deviations, and slopes for Ray's daily time-series phase data.

DV	Baseline (N=30)			TA (N=45)			Follow-up (N=26)			Total (N=101)		
	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope
1.	3.56	1.10	.004	3.42	1.09	-.03	2.88	.58	-.03	3.32	1.02	-.01
2.	3.65	.96	.01	3.55	1.15	-.04	3.04	.74	-.01	3.45	1.03	-.01
3.	3.64	.63	.001	2.62	.80	-.02	2.32	.38	.002	2.53	.68	-.005
4.	3.28	1.83	.13	3.06	1.52	-.02	2.98	1.39	-.05	3.11	1.59	-.003
5.	2.33	.95	-.002	1.87	1.04	-.02	1.39	.41	-.01	1.88	.96	-.01
6.	3.42	1.44	.02	3.23	1.20	-.03	2.90	.87	-.02	3.20	1.22	-.01
7.	4.64	1.67	-.09	4.51	1.86	.03	4.84	1.52	.05	4.63	1.73	.003

Note: **DV** = Dependent variable; **1** = General feelings toward Mary (1=Extremely positive, 7=Extremely negative); **2** = Feeling of closeness to Mary (1=Extremely close, 7=Extremely distant); **3** = Hopefulness about relationship (1=very hopeful, 5=not at all hopeful); **4** = Degree to which Ray felt bothered by Mary acting demanding or jealous (1=not at all bothered, 7=extremely bothered); **5** = Degree to which Ray was open vs. private with Mary (1=very open, 7=very private); **6** = Degree of tension with Mary (1=no tension, 7=extreme tension); **7** = Degree of explosiveness in relationship (1=extremely explosive, 7=not at all explosive).

Table A2. Means, standard deviations, and slopes for Mary's daily time-series phase data.

DV	Baseline (N=30)			TA (N=45)			Follow-up (N=27)			Total (N=102)		
	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope
1.	3.26	.49	.01	3.74	1.08	-.03	3.25	.87	-.03	3.47	.92	-.002
2.	3.44	.75	.03	3.79	1.05	-.03	3.60	.94	-.04	3.64	.95	.00
3.	2.36	.46	.03	2.59	.80	-.03	2.22	.48	-.01	2.42	.65	-.003
4.	3.40	1.01	.07	3.81	1.12	-.02	3.34	.89	-.02	3.57	1.06	.00
5.	3.59	.98	.03	3.97	1.26	-.02	3.80	1.19	-.05	3.81	1.17	.00
6.	3.57	1.25	-.02	3.61	1.45	-.07	3.22	1.16	-.04	3.49	1.33	-.01
7.	2.61	.65	.01	2.77	1.41	-.04	2.49	1.28	-.07	2.65	1.20	-.01
8.	5.96	1.02	-.03	5.95	1.48	.01	5.96	1.60	.08	5.96	1.39	.001

Note: **DV** = Dependent variable; **1** = General feelings toward Ray (1=Extremely positive, 7=Extremely negative); **2** = Feeling of closeness to Ray (1=Extremely close, 7=Extremely distant); **3** = Hopefulness about relationship (1=very hopeful, 5=not at all hopeful); **4** = Degree to which Ray was open vs. private with Mary (1=very open, 7=very private); **5** = Degree to which Mary felt valued/loved (1=extremely valued/loved, 7=not at all valued/loved); **6** = Feelings of self-worth (1=a great deal of self-worth, 7=no self-worth); **7** = Degree of tension with Ray (1=no tension, 7=extreme tension); **8** = Degree of explosiveness in relationship (1=extremely explosive, 7=not at all explosive).

Table A3. Means, standard deviations, and slopes for George's daily time-series phase data.

DV	Baseline (N=28)			TA (N=39)			Follow-up (N=28)			Total (N=95)		
	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope
1.	2.48	.54	-.01	2.20	.38	.01	2.49	.67	-.001	2.37	.55	.00
2.	2.68	.63	-.02	2.36	.47	.004	2.76	.77	.01	2.57	.64	.00
3.	1.97	.54	-.01	1.39	.48	-.02	1.68	.64	.02	1.65	.60	-.01
4.	2.25	1.11	-.02	2.74	1.12	.02	2.02	1.39	-.02	2.38	1.24	-.003
5.	3.02	.91	-.02	2.50	.54	-.01	2.53	.55	.002	2.66	.71	-.01
6.	2.90	.96	-.002	2.90	.50	.01	3.26	.68	-.01	3.01	.73	.01

Note: **DV** = Dependent variable; **1** = General feelings toward Susan (1=Extremely positive, 7=Extremely negative); **2** = Feeling of closeness to Susan (1=Extremely close, 7=Extremely distant); **3** = Hopefulness about relationship (1=very hopeful, 5=not at all hopeful); **4** = Degree to which George and Susan got caught up in a “power struggle” (1= not at all, 7=extremely); **5** = Degree to which George felt loved (1=extremely loved, 7=not at all loved); **6** = Degree to which Susan was emotionally honest/open with George (1= extremely open; 7=not at all open).

Table A4. Means, standard deviations, and slopes for Susan’s daily time-series phase data.

DV	Baseline (N=28)			TA (N=39)			Follow-up (N=27)			Total (N=94)		
	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope
1.	2.21	.67	.02	2.43	.78	-.02	2.46	.83	-.02	2.38	.77	.002
2.	2.54	.63	.01	2.67	.73	-.02	2.85	.97	-.01	2.68	.79	.002
3.	2.11	.49	.02	2.21	.46	-.004	2.31	.65	-.04	2.21	.54	.002
4.	1.50	.63	-.01	1.59	.67	-.02	1.54	.95	-.02	1.55	.75	-.002
5.	2.36	.67	.01	2.77	.83	-.02	2.58	.82	-.01	2.59	.80	.002
6.	1.46	.50	-.02	1.56	.81	-.01	1.16	.45	-.01	1.42	.66	-.01
7.	6.57	.78	.004	6.54	.81	.03	6.65	1.05	.03	6.59	.87	.003

Note: **DV** = Dependent variable; **1** = General feelings toward George (1=Extremely positive, 7=Extremely negative); **2** = Feeling of closeness to George (1=Extremely close, 7=Extremely distant); **3** = Hopefulness about relationship (1=very hopeful, 5=not at all hopeful); **4** = Degree to which Susan felt her needs and wants were not as important as George’s (1= not at all, 7=extremely); **5** = Degree to which Susan felt loved (1=extremely loved, 7=not at all loved); **6** = Degree to which Susan was bothered by George correcting or challenging her to defend her opinions (1=not at all bothered, 7=extremely bothered); **7** = Anger toward George (1=Extremely angry, 7=not at all angry).

Table A5. Means, standard deviations, and slopes for Tom's daily time-series phase data.

DV	Baseline (N=29)			TA (N=53)			Follow-up (N=28)			Total (N=110)		
	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope
1.	2.03	.18	-.01	1.97	.27	.002	2.11	.41	-.01	2.02	.30	.00
2.	2.83	.68	-.002	2.23	.60	.01	2.15	.44	-.01	2.37	.65	-.01
3.	1.49	.47	-.02	1.09	.23	.002	1.11	.31	-.01	1.20	.37	-.004
4.	4.37	1.26	-.02	4.14	1.55	.02	3.12	1.13	-.06	3.94	1.46	-.01
5.	2.07	.25	-.01	1.95	.23	.003	2.11	.31	-.01	2.02	.27	.00
6.	2.32	1.08	-.002	2.11	.84	-.03	3.07	1.57	.02	2.62	1.36	.00

Note: **DV** = Dependent variable; **1** = General feelings toward Elaine (1=Extremely positive, 7=Extremely negative); **2** = Feeling of closeness to Elaine (1=Extremely close, 7=Extremely distant); **3** = Hopefulness about relationship (1=very hopeful, 5=not at all hopeful); **4** = Degree to which Tom's attention was focused on the relationship (1=focused almost exclusively on the relationship, 7=almost exclusively focused on other things); **5** = Degree of confidence in relationship (1=extremely confident, 7=not at all confident); **6** = Degree of tension/irritability (1=not at all tense or irritable, 7= extremely tense or irritable).

Table A6. Means, standard deviations, and slopes for Elaine's daily time-series phase data.

DV	Baseline (N=29)			TA (N=53)			Follow-up (N=35)			Total (N=117)		
	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope	Mean	SD	Slope
1.	1.65	.45	.02	1.50	.47	-.02	1.11	.28	.01	1.42	.47	-.01
2.	1.80	.37	.01	1.91	.26	-.001	1.90	.28	-.01	1.88	.30	.00
3.	1.00	.00	.00	1.00	.00	.00	1.00	.00	.00	1.00	.00	.00
4.	2.66	.69	-.02	2.58	.90	.02	2.15	.67	-.02	2.47	.82	-.004
5.	1.74	.50	-.01	1.64	.45	-.02	1.05	.17	.003	1.49	.49	-.01
6.	2.80	.79	-.05	2.30	.72	.01	2.29	.76	-.03	2.29	.75	-.01

Note: **DV** = Dependent variable; **1** = General feelings toward Tom (1=Extremely positive, 7=Extremely negative); **2** = Feeling of closeness to Tom (1=Extremely close, 7=Extremely distant); **3** = Hopefulness about relationship (1=very hopeful, 5=not at all hopeful); **4** = Degree to which Elaine's attention was focused on the relationship (1=focused almost exclusively on the relationship, 7=almost exclusively focused on other things); **5** = Degree of confidence in relationship (1=extremely confident, 7=not at all confident); **6** = Anxiety (1=not at all anxious; 7=extremely anxious).

Appendix K: SEQ-C means and standard deviations

Table A7. *SEQ-C means and standard deviations by phase: Mary and Ray*

	Baseline (n=2)		TA (n=3)		Follow-up (n=4)	
Mary	Mean	SD	Mean	SD	Mean	SD
Depth	4.40	.57	5.80	.53	4.45	1.23
Smoothness	2.50	.71	2.07	1.17	3.85	2.39
Positivity	2.60	.85	2.67	.50	4.05	2.19
Arousal	3.60	.57	2.67	.23	3.50	1.05
	Baseline (n=2)		TA (n=4)		Follow-up (n=1)	
Ray	Mean	SD	Mean	SD	Mean	SD
Depth	5.30	.99	5.95	.44	6.20	n/a
Smoothness	2.60	1.98	3.15	.93	2.20	n/a
Positivity	3.40	1.98	3.60	.67	3.60	n/a
Arousal	3.90	.71	3.85	.53	4.40	n/a

Note: n=number of measurements per phase.

Table A8. *SEQ-C means and standard deviations by phase: Susan and George*

	Baseline (n=3)		TA (n=6)		Follow-up (n=3)	
Susan	Mean	SD	Mean	SD	Mean	SD
Depth	4.47	.99	4.40	.81	5.07	.83
Smoothness	4.00	.87	4.43	.89	3.60	1.40
Positivity	4.20	1.11	4.50	.88	4.07	1.11
Arousal	3.46	.12	3.60	.96	4.40	.53
	Baseline (n=3)		TA (n=6)		Follow-up (n=3)	
George	Mean	SD	Mean	SD	Mean	SD
Depth	5.40	1.00	5.07	.92	5.73	.58
Smoothness	3.13	1.01	3.90	1.14	3.13	.61
Positivity	4.20	1.40	4.77	1.01	4.80	.60
Arousal	4.00	1.22	4.23	.69	4.73	.12

Note: n=number of measurements per phase.

Table A9. *SEQ means and standard deviations by phase: Elaine and Tom*

	Baseline (n=5)		TA (n=3)		Follow-up (n=3)	
Elaine	Mean	SD	Mean	SD	Mean	SD
Depth	4.92	.58	5.27	.58	4.80	.87
Smoothness	3.92	1.16	4.07	.90	3.40	1.11
Positivity	3.96	.95	4.27	.70	3.87	1.03
Arousal	4.04	.36	4.07	.12	3.93	.12
	Baseline (n=5)		TA (n=3)		Follow-up (n=2)	
Tom	Mean	SD	Mean	SD	Mean	SD
Depth	4.08	1.15	4.53	1.03	5.10	.14
Smoothness	4.00	1.10	4.33	.99	2.70	.42
Positivity	4.08	.52	4.53	.23	3.10	.14
Arousal	3.52	1.01	4.40	1.04	3.40	0.00

Note: n=number of measurements per phase.

Appendix L: Tables of BSI, MDS, and DAS scores

Table A10. *Marital satisfaction and psychological symptoms: Mary and Ray*

Mary	BSI (raw score)	BSI (T-score)	MDS	DAS
Time 1	50	64	33	104
Time 2	29	59	30	110
Time 3	6	43	25	118
Ray*				
Time 1	18	57	55	88
Time 3	26	61	54	97

*Ray did not complete measures at Time 2.

Table A11. *Marital satisfaction and psychological symptoms: Susan and George*

Susan	BSI (raw scores)	BSI (T-scores)	MDS	DAS
Time 1	55	66	44	93
Time 2	32	60	39	95
Time 3	32	60	35	107
George				
Time 1	40	67	32	109
Time 2	40	66	31	120
Time 3	52	72	32	112

Table A12. *Marital satisfaction and psychological symptoms: Elaine and Tom*

Elaine	BSI (raw scores)	BSI (T-scores)	MDS	DAS
Time 1	34	61	23	121
Time 2	14	51	25	123
Time 3	21	56	21	119
Tom				
Time 1	20	58	30	114
Time 2	24	60	22	123
Time 3	10	50	23	119

Appendix M: SEQ-T Tables

Table A13. *SEQ-T means and standard deviations by phase: Mary and Ray's therapist*

		Mean	SD
Baseline n=2	Depth	4.20	.57
	Smoothness	3.20	.85
TA n=4	Depth	5.40	1.18
	Smoothness	1.95	.81
Follow-up n=4	Depth	4.05	.91
	Smoothness	3.85	1.58

Table A14. *SEQ-T means by phase: Susan and George's therapist*

		Mean	SD
Baseline n=3	Depth	4.67	.31
	Smoothness	3.80	.92
TA n=5	Depth	5.16	.89
	Smoothness	3.84	1.01
Follow-up n=4	Depth	5.05	.41
	Smoothness	3.45	1.22

Table A15. *SEQ-T means by phase: Elaine and Tom's therapist*

		Mean	SD
Baseline n=5	Depth	4.36	1.13
	Smoothness	3.84	.17
TA n=3	Depth	5.13	1.10
	Smoothness	3.93	.12
Follow-up n=3	Depth	4.27	.31
	Smoothness	3.87	.23

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*(1), 82-109.
- Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling, 25*, 165-181.
- American Psychological Association (2002). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author.
- APA Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist, 61*, 271-285.
- Aranow, E., Reznikoff, M., & Moreland, K. (1994). *The Rorschach technique: Perceptual basics, content interpretation, and applications*. Boston: Allyn and Bacon.
- Bender, L. A. (1938). *Visual Motor Gestalt Test and its clinical use: Research monograph number 3*. New York: American Orthopsychiatric Association.
- Borckardt, J. J. (2008). *Simulation Modeling Analysis Version 8.3.3 User's Guide*. Jeffery Borckardt.
- 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: A guide to case-based time series analysis. *American Psychologist, 63*(2), 77-95.
- Bruhn, A. R. (1992a). *Memories of spouse*. Bethesda, MD: Arnold R. Bruhn and Associates.
- Bruhn, A. R. (1992b). The early memories procedure: A projective test of autobiographical memory, Part 1. *Journal of Personality Assessment, 58*(1), 1-15.
- Bruhn, A. R. (1992c). The early memories procedure: A projective test of autobiographical memory, Part 2. *Journal of Personality Assessment, 58*(2), 326-346.
- Butcher, J. N. (1990). *The MMPI-2 in psychological treatment*. New York: Oxford University Press.

- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.
- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *Minnesota Multiphasic Personality Inventory-2: Manual for administration, scoring and interpretation* (Rev. ed.). Minneapolis: University of Minnesota Press.
- Butcher, J. N., & Williams, C. L. (2000). *Essentials of MMPI-2 and MMPI-A interpretation* (2nd ed.). Minneapolis: University of Minnesota Press.
- Chambless, D. L., & Hollon, S. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7-18.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685-716.
- Crosbie, J. (1993). Interrupted time-series analysis with brief single-subject data. *Journal of Consulting and Clinical Psychology, 61*(6), 966-974.
- Cummings, E., & Davies, P. (2002). Effects of marital conflict on children: Recent advances and emerging themes in process-oriented research. *Journal of Child Psychology and Psychiatry, 43*(1), 31-63.
- Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*(1), 1-38.
- Derogatis, L. R. (1993). *Brief Symptom Inventory (BSI): Administration, scoring, and procedures manual*. Baltimore: Clinical Psychometric Research.
- Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13*(3), 595-605.
- Derogatis, L. R., Rickels, K., & Rock, A. F. (1976). The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *British Journal of Psychiatry, 128*, 280-289.
- Dorr, D. (1981). Conjoint psychological testing in marriage therapy: New wine in old skins. *Professional Psychology, 12*(5), 549-555.

- Doss, B. D., Simpson, L. E., & Christensen, A. (2004). Why do couples seek marital therapy? *Professional Psychology: Research and Practice*, 35(6), 608-614.
- Exner, J. E., Jr. (1974). *The Rorschach: A comprehensive system: Vol. 1*. New York: Wiley.
- Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations (3rd ed.)*. New York: Wiley.
- Finn, S. E. (1996). *Manual for using the MMPI-2 as a therapeutic intervention*. Minneapolis: University of Minnesota Press.
- Finn, S. E. (2007). *In our clients' shoes: Theory and techniques of Therapeutic Assessment*. Mahwah, NJ: Erlbaum.
- Finn, S. E., Schroeder, D. G., & Tonsager, M. E. (1994). *The Assessment Questionnaire-2 (AQ-2): A measure of clients' experiences with psychological assessment*. Unpublished manuscript.
- Finn, S. E., & Tonsager, M. E. (1992). The therapeutic effects of providing MMPI-2 test feedback to college students awaiting psychotherapy. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 4, 278-287.
- Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment*, 9(4), 374-385.
- Fischer, C. T. (1970). The testee as co-evaluator. *Journal of Counseling Psychology*, 17(1), 70-76.
- Fischer, C. T. (1972). Paradigm changes which allow sharing of results. *Professional Psychology*, 3, 364-369.
- Fischer, C. T. (1994). *Individualizing psychological assessment*. Hillsdale, NJ: Erlbaum (Original work published 1985).
- Floyd, F., Haynes, S., & Kelly, S. (1997). Marital assessment: A dynamic functional-analytic approach. *Clinical handbook of marriage and couples interventions* (pp. 349-377). Hoboken, NJ, US: John Wiley & Sons Inc.
- Ganellen, R. J. (1996). Comparing the diagnostic efficiency of the MMPI, MCMI-II, and Rorschach: A review. *Journal of Personality Assessment*, 67(2), 219-243.

- Graham, J., Liu, Y., & Jeziorski, J. (2006). The Dyadic Adjustment Scale: A Reliability Generalization Meta-Analysis. *Journal of Marriage and Family*, 68(3), 701-717.
- Greenberg, L. S., & Johnson, S. M. (1988). *Emotionally focused therapy for couples*. New York: Guilford Press.
- Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th ed.). Hoboken, NJ: John Wiley & Sons.
- Handler, L. (1997). He says, she says, they say: The consensus Rorschach. In J. R. Meloy, M. W. Acklin, C. B. Gacono, J. F. Murray, & C. A. Peterson (Eds.), *Contemporary Rorschach interpretation* (pp. 499-533). Mahwah, NJ: Erlbaum.
- Hanson, W., Claiborn, C., & Kerr, B. (1997). Differential effects of two test-interpretation styles in counseling: A field study. *Journal of Counseling Psychology*, 44(4), 400-405.
- Heppner, P. P., Kivlighan, D. M., & Wampold, B. E. (1999). *Research design in counseling* (2nd ed.). Belmont, CA: Wadsworth.
- Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (1999). A comparative analysis of Rorschach and MMPI validity. *Psychological Assessment*, 11(3), 278-296.
- Hilsenroth, M. J., Peters, E. J., & Ackerman, S. J. (2004). The development of therapeutic alliance during psychology assessment: Patient and therapist perspectives across treatment. *Journal of Personality Assessment*, 83, 331-344.
- Jacobson, N. S. (1981). Behavioral marital therapy. In A. S. Gurman & D. P. Kniskern (Eds.) *Handbook of family therapy*. (pp. 556-591). New York: Brunner/Mazel.
- Jacobson, N. S., & Margolin, G. (1979). *Marital therapy: Strategies based on social learning and behavior exchange principles*. New York: Brunner/Mazel.
- Johnson, S. M. (2004). *The practice of emotionally focused couple therapy: Creating connection*. New York: Brunner-Routledge.
- Johnson, S. M., Makinen, J. A., & Millikin, J. W. (2001). Attachment injuries in couple relationships: A new perspective on impasses in couples therapy. *Journal of Marital and Family Therapy*, 27(2), 145-155.
- Kayser, K. (1996). The Marital Disaffection Scale: An inventory for assessing emotional estrangement in marriage. *American Journal of Family Therapy*, 24(1), 83-88.

- Kazdin, A. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- 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(404), 1198-1202.
- Locke, H., & Wallace, K. (1959). Short marital-adjustment and prediction tests: Their reliability and validity. *Marriage & Family Living*, 21, 251-255.
- Meyer, G. J., Hilsenroth, M. J., Baxter, D., Exner, J. E., Fowler, J. C., Piers, C. C., et al. (2002). An examination of interrater reliability for scoring the Rorschach Comprehensive System in eight data sets. *Journal of Personality Assessment*, 78, 219–274.
- Middelberg, C. V. (2001). Projective identification in common couple dances. *Journal of Marital and Family Therapy*, 27(3), 341-352.
- Millon, T., Weiss, L., Millon, C., & Davis, R. D. (1994). *MIPS: Millon Index of Personality Styles Manual*, San Antonio: Psychological Corp.
- Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI-2 test feedback to clients at a university counseling service. *Psychological Assessment*, 9(2), 122-131.
- Noy-Sharav, D. (2005). The Rorschach and the TAT as relational instruments: Evaluating young couples with Consensus Rorschach and TAT. *Rorschachiana*, 27, 139-163.
- Nurse, A. R. (1999). *Family assessment: Effective uses of personality tests with couples and families*. New York: John Wiley & Sons, Inc.
- Priest, R., & Thein, M. (2003). Humor appreciation in marriage: Spousal similarity, associative mating, and disaffection. *Humor: International Journal of Humor Research*, 16(1), 63-78.
- Richman, J., & Davidoff, I. (1971). Interaction testing and counseling as a form of crisis intervention during marital therapy. *Proceedings of the Annual Convention of the American Psychological Association*, 6(1), 439-440.
- Rorschach, H. (1921/1942). *Psychodiagnostics*. Bern: Allyn & Bacon.
- Sabourin, S., Valois, P., & Lussier, Y. (2005). Development and validation of a brief version of the Dyadic Adjustment Scale with a nonparametric item analysis model. *Psychological Assessment*, 17, 15 – 27.

- Schramm, D. G. (2006). Individual and social costs of divorce in Utah. *Journal of Family and Economic Issues*, 27(1), 133-151.
- Schroeder, D. G., Hahn, E. D., Finn, S. E., & Swann, W. B. (1993). *Personality feedback has more impact when mildly discrepant from self-views*. Paper presented at the fifth annual convention of the American Psychological Society, Chicago, IL.
- Shadish, W. R., & Baldwin, S. (2003). Meta-analysis of MFT interventions. *Journal of Marital and Family Therapy*, 29(4), 547-570.
- Shadish, W. R., & Baldwin, S. (2005). Effects of Behavioral Marital Therapy: A meta-analysis of randomized controlled trials. *Journal of Consulting and Clinical Psychology*, 73(1), 6-14.
- Shadish, W. R., Ragsdale, K., Glaser, R. R., & Montgomery, L. M. (1995). The efficacy and effectiveness of marital and family therapy: A perspective from meta-analysis. *Journal of Marital and Family Therapy*, 21(4), 345-360.
- Skinner, H., Steinhauer, P., & Santa-Barbara, J. (1983). The Family Assessment Measure. *Canadian Journal of Community Mental Health*, 2(2), 91-105.
- Skinner, H., Steinhauer, P., & Sitarenios, G. (2000). Family Assessment Measure (FAM) and Process Model of Family Functioning. *Journal of Family Therapy*, 22(2), 190-210.
- Smith, J. D., Wolf, N. J., Handler, L., & Nash, M. R. (2009). Testing the effectiveness of family Therapeutic Assessment: A case study using a time-series design. *Journal of Personality Assessment*, 91(6), 518-533.
- Smith, S. R., Wiggins, C. M., & Gorske, T. T. (2007). A survey of psychological assessment feedback practices. *Assessment*, 14(3), 310-319.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, 38(1), 15-28.
- Steinhauer, P., Santa-Barbara, J., & Skinner, H. (1984). The process model of family functioning. *The Canadian Journal of Psychiatry / La Revue canadienne de psychiatrie*, 29(2), 77-88.
- Stewart, R., & Chambless, D. (2008). Treatment failures in private practice: How do psychologists proceed? *Professional Psychology: Research and Practice*, 39(2), 176-181.

- Stiles, W. B., Gordon, L. E., & Lani, J. A. (2002). Session evaluation and the Session Evaluation Questionnaire. In G. S. Tryon (Ed.), *Counseling based on process research: Applying what we know* (pp. 325-343). Boston, MA: Allyn & Bacon.
- Stiles, W. B., Reynolds, S., Hardy, G. E., Rees, A., Barkham, M., & Shapiro, D. A. (1994). Evaluation of description of psychotherapy sessions by clients using the Session Evaluation Questionnaire and the Session Impacts Scale. *Journal of Counseling Psychology, 41*(2), 175–185.
- Swann, W. B. (1997). The trouble with change: Self-verification and allegiance to the self. *Psychological Science, 8*, 177-180.
- Tharinger, D., Finn, S. E., Gentry, L., Hamilton, A., Fowler, J., Matson, M., Krumholz, L., & Walkowiak, J. (2009). Therapeutic Assessment with children: A pilot study of treatment acceptability and outcome. *Journal of Personality Assessment 91*(3), 238-244.
- Tharinger, D., Finn, S. E., Wilkinson, A., & Schaber, P. (2007). Therapeutic assessment with a child as a family intervention: A clinical and research case study. *Psychology in the Schools, 44*(3), 293-309.
- Velicer, W. F., & Colby, M. (2005). A comparison of missing-data procedures for ARIMA time-series analysis. *Educational and Psychological Measurement, 65*(4), 596-615.
- Weiss, R. L. & Tolman, A. O. (1990). The Marital Interaction Coding System—Global (MICS-G): A global companion to the MICS. *Behavioral Assessment, 12*, 271-294.
- Whisman, M. (2007). Marital distress and DSM-IV psychiatric disorders in a population-based national survey. *Journal of Abnormal Psychology, 116*(3), 638-643.
- Whisman, M., & Bruce, M. (1999). Marital dissatisfaction and incidence of major depressive episode in a community sample. *Journal of Abnormal Psychology, 108*(4), 674-678.
- Whisman, M., Uebelacker, L., & Bruce, M. (2006). Longitudinal association between marital dissatisfaction and alcohol use disorders in a community sample. *Journal of Family Psychology, 20*(1), 164-167.
- Wood, N. D., Crane, D. R., Schaalje, G. B., & Law, D. D. (2005). What works for whom: A meta-analytic review of marital and couples therapy in reference to marital distress. *The American Journal of Family Therapy, 33*, 273-287.

Worthington, E. L., McCullough, M. E., Shortz, J. L., Mindes, E. J., Sandage, S. J., & Chartrand, J. M. (1995). Can couples assessment and feedback improve relationships? Assessment as a brief relationship enrichment procedure. *Journal of Counseling Psychology*, 42(4), 466-475.

VITA

Jennifer Anne Durham-Fowler attended Falls Church High School in Falls Church, Virginia. In 1997 she entered Pepperdine University in Malibu, California. She received the degree of Bachelor of Arts from Pepperdine University in April, 2001. During the following years she was employed as a research and teaching assistant in the psychology department at the University of California, Los Angeles. In September, 2004, she entered the Counseling Psychology doctoral program at The University of Texas at Austin. In 2008 she received her master's degree in Educational Psychology from The University of Texas at Austin. She is currently in the process of completing her pre-doctoral internship at the Danielsen Institute at Boston University.

Permanent Address: 24 Highland Avenue #21, Cambridge, MA 02139.

This manuscript was typed by the author.