

Measuring the Therapist's Impact on the Patient's Therapeutic Progress

George Silberschatz and John T. Curtis

This article describes methods and concepts developed by the Mount Zion Psychotherapy Research Group for empirically evaluating the pertinence or suitability of a therapist's interventions (behaviors) to a patient's particular problems, needs, and treatment goals. Intensive studies of 2 brief psychotherapy cases are presented. In these studies, patient-initiated critical incidents (tests) were identified, the case-specific accuracy of the therapist's responses to these incidents was rated, and the impact of these interventions on subsequent patient behavior was measured. The findings indicated that these patients tended to show improvement in the therapeutic process when the therapist's interventions were in accord with their particular problems and treatment goals. The application of this method to clinically relevant studies of psychotherapy is discussed.

Inherent in most theories of psychotherapy is the assumption that the psychotherapist's attitudes and technical interventions play a critical part in facilitating (or sometimes impeding) a patient's therapeutic progress. Despite voluminous research on the therapist's contribution to psychotherapy (for comprehensive reviews, see Lambert & Bergin, 1983; Lambert, Shapiro, & Bergin, 1986; Parloff, Waskow, & Wolfe, 1978; Schaffer, 1982), surprisingly little progress has been made in understanding or empirically demonstrating how the therapist contributes to the success or failure of psychotherapy. Many reviewers have suggested that this lack of progress is due to (a) inadequate conceptualization of how therapist interventions affect particular patients and (b) methods of evaluating therapist behaviors that are too global, imprecise, or clinically irrelevant (Fiske, 1977; Greenberg & Pinsof, 1986; Jones, Cummings, & Horowitz, 1988; Persons, 1991; Stiles, Shapiro, & Elliott, 1986; Strupp, 1986).

One major problem in much of the literature is that investigators have attempted to assess therapist interventions without regard to a particular patient's specific problems and needs (Rice & Greenberg, 1984; Silberschatz, Curtis, & Nathans, 1989; Stiles et al., 1986). In response to this issue, several recent studies have developed case-specific measures that assess the accuracy of therapist interpretations (Crits-Christoph, Cooper, & Luborsky, 1988; Silberschatz, Fretter, & Curtis, 1986). In this article, we describe methods and concepts developed by the Mount Zion Psychotherapy Research Group for empirically

evaluating the pertinence or suitability of a therapist's interventions (behaviors) to a patient's particular problems, needs, and treatment goals.

Our evaluation of the suitability of a therapist's interventions of course reflects the theory of psychotherapy we follow. We therefore present a brief overview of this model of therapy—the control-mastery theory—a contemporary psychoanalytic theory developed by Weiss (1986). This theory offers a precise and explicit model of how a therapist's attitudes and behavior may influence the progress of a particular patient with his or her specific problems, goals, and ways of working in therapy. Finally, we present a study in which we empirically tested specific predictions based on the model.

BACKGROUND

Overview of Our Theoretical Perspective

The control-mastery theory is a cognitive psychoanalytic model of psychopathology and psychotherapy (Weiss, 1986). According to the theory, human behavior is regulated by ideas, and psychopathology stems from unconscious false beliefs—"pathogenic beliefs"—that are based on traumatic childhood experiences. A pathogenic belief may be viewed as a compelling unconscious theory formed in early childhood that warns the person who adheres to it against pursuing a particular goal. Thus, pathogenic beliefs are the primary sources of resistance to treatment. Nonetheless, patients are motivated to overcome pathogenic beliefs because they are grim, constricting, and lead to renunciation of important life goals.

In psychotherapy, the patient works to disconfirm pathogenic beliefs by *testing* them in the relationship with the therapist. Indeed, testing these beliefs is one of the patient's central activities. In testing a pathogenic belief, the patient plans and carries out a trial action that is intended to provide information about the belief and about the patient's effect on the therapist. One way in which a patient might test a pathogenic belief is to behave in ways associated with the advent of trauma. For example, a patient who experienced rejection after making demands on her parents might act in a demanding or complaining fashion with the therapist to test whether such behavior results in censure. This style of testing is referred to as transference test-

George Silberschatz and John T. Curtis, Department of Psychiatry, Mount Zion Medical Center of the University of California, San Francisco.

Preparation of this article was supported in part by National Institute of Mental Health Grant MH35230 and by the Research Support Program of Mount Zion Hospital and Medical Center.

The invaluable support of the San Francisco Psychotherapy Research Group for this research is gratefully acknowledged. Thomas J. Kelly and Polly B. Fretter made particular contributions to the collection and analysis of data.

Correspondence concerning this article should be addressed to George Silberschatz, 3368 Sacramento Street, San Francisco, California 94118-1912.

ing (an unfortunate name because all testing occurs in the transference).

Another broad type of testing involves the patient behaving in a trauma-inducing way toward the therapist. This form of testing, called passive-into-active testing, was illustrated by a patient who experienced repeated criticisms and condemnations from parents. In therapy, he was critical and condemning of the therapist, both to ascertain if the therapist was able to handle this behavior and—by observing the therapist's reaction—to learn more adaptive responses. According to the theory, the patient's primary purpose is to disconfirm false beliefs; consequently, the patient hopes that the therapist will not react as the parents once reacted. If the therapist does not, the patient is reassured, and the strength of the pathogenic belief is lessened.

For example, a patient remembered that her mother became depressed in reaction to the patient's childhood attempts at separation and autonomy. Whenever she left her mother to go to school or play with friends, her mother became withdrawn, sullen, or depressed; when she stayed home, her mother seemed happy. This patient developed the pathogenic belief that her autonomy was harmful or upsetting to others and consequently stifled important needs and goals. The patient worked to disconfirm this pathogenic belief by behaving independently in therapy (e.g., by disagreeing with the therapist, coming up with her own insights, and arriving late for sessions) to see if the therapist (unlike her mother) could tolerate her autonomous strivings. When the therapist did not become defensive or act critically toward the patient, he passed this test; that is, the therapist's behavior disconfirmed the patient's pathogenic belief that her autonomy would hurt the therapist. The therapist would have failed this test by acting in a way that the patient would have experienced as being hurt by or critical of her independence. Such a response would have confirmed the patient's pathogenic expectation that it was dangerous to behave autonomously and thus would have been countertherapeutic.

Each patient has specific pathogenic beliefs stemming from particular traumatic life experiences. Although there may be some similarity across patients, it does not follow that patients who fit into similar diagnostic categories will necessarily have similar beliefs, methods of testing, or therapy goals. The control-mastery theory does not propose new techniques for dealing with a particular category of patients; rather, it offers a model for conceptualizing how a person's problems arose and the ways in which the patient will need to test the therapist to disconfirm pathogenic beliefs. Tests may vary according to their relevance to pathogenic beliefs; *key tests* are those that are most critical to the patient because they are central to the predominant pathogenic beliefs that the patient is working to disconfirm (for a thorough exposition of the concept of testing, see Curtis & Silberschatz, 1986; Silberschatz & Curtis, 1986; Weiss, 1986).

Review of Previous Research

Previous empirical studies have demonstrated that pathogenic beliefs, and the ways in which patients are likely to test therapists to disconfirm them, can be reliably inferred and formulated from early therapy sessions (Curtis, Silberschatz,

Sampson, & Weiss, in press; Curtis, Silberschatz, Sampson, Weiss, & Rosenberg, 1988; Rosenberg, Silberschatz, Curtis, Sampson, & Weiss, 1986; see also Collins & Messer, 1988, 1991; Perry, Luborsky, Silberschatz, & Popp, 1989). Research has also demonstrated that trained clinicians can make reliable judgments of whether a therapist has passed or failed a patient's tests and that ratings of the degree to which a test is passed or failed correlate significantly with immediate therapeutic progress. For instance, Horowitz, Sampson, Siegelman, Wolfson, and Weiss (1975) found that a patient's level of anxiety consistently dropped and that new contents (previously warded off) tended to emerge when the therapist passed a test. Silberschatz (1978, 1986) identified 46 key tests in the verbatim transcripts of the first 100 hr of a psychoanalysis. Correlations between ratings of the degree to which the therapist passed or failed these key tests and changes in a variety of patient measures indicated that the patient became significantly more involved, more productive, and more relaxed when the therapist passed a key test. Research has also shown that there is a significant correlation between the "plan compatibility" of therapist interpretations (another measure of the degree to which an intervention confirms or disconfirms a patient's pathogenic beliefs) and immediate patient therapeutic progress (Silberschatz, Fretter, & Curtis, 1986).

The research just cited involved the use of repeated measures and intensive single-case designs. These studies followed a broad research strategy described as the "events paradigm" in psychotherapy process research (Stiles et al., 1986). The events approach asks, "Which specific therapist interventions introduced in which momentary contexts will lead to which immediate and subsequent impacts . . . ?" (p. 174). We have adopted the events paradigm in our studies of psychotherapy process because we believe that it most closely reflects the actual dilemmas that therapists face, and as a result, findings from such research are most likely to be of relevance and value to clinical practice. We further believe that at this stage of understanding therapeutic process, the events paradigm is best performed in single case studies. Only in such studies can the individual meaning of a single event be adequately described and investigated. However, there are obvious limitations to single-case designs; findings need to be replicated on other cases and with other measures.

In this study, we intend to contribute to this pool of data by proposing that a therapist's disconfirmation of the patient's pathogenic belief in the context of a patient testing a central pathogenic belief (key test) will lead to immediate patient productivity. Unlike our earlier research on testing, in this study we focus on brief psychodynamic psychotherapy rather than on psychoanalysis.

METHOD

Subjects

Patients

Data for the study were obtained from the Mount Zion Brief Therapy Research Project, a study investigating the process and outcome of brief (16-session) psychodynamic therapy. All patients were self-re-

ferred and were screened to ensure their suitability for brief treatment. The screening included a pretherapy evaluation questionnaire and a clinical intake interview with an independent evaluator. Acceptance for brief therapy was based on the clinical assessment and the questionnaire data. Patients were required to meet the following minimum acceptance criteria: (a) a history of positive interpersonal relationships; (b) no evidence of psychosis, organic brain syndrome, or mental retardation; (c) no evidence of serious substance abuse; and (d) no evidence of suicidal potential. After the intake interview, patients who met these criteria were asked to sign an informed consent form, which explained that they would be interviewed by an independent evaluator immediately after therapy and at 6-month and 1-year follow-up sessions, at which times they would also be asked to fill out forms and questionnaires. They were also informed that all interviews and therapy sessions would be audio recorded, transcribed, and used for research purposes. Strict anonymity and confidentiality were guaranteed by a systematic coding process that deleted all names and other potential identifying data (all names and information presented in this article have been changed to protect confidentiality).

For our study, two patients—Diane, a 34-year-old divorced attorney, and Gary, a 36-year-old married graduate student—were randomly selected from the larger sample of patients. Although the patients had different problems and backgrounds, they were very similar with respect to the general nature and severity of their psychopathology and to the general nature of their pathogenic beliefs (as described later). On the basis of the intake interviews, self-report measures, and independent clinical evaluator measures, both patients were diagnosed as suffering from a dysthymic disorder (*Diagnostic and Statistical Manual of Mental Disorders*, 3rd ed., rev; American Psychiatric Association, 1987). Diane completed 16 therapy sessions, and Gary completed 12. At the conclusion of therapy, both patients showed excellent improvement based on ratings and evaluations made by the patient, the therapist, and an independent clinical evaluator.

Therapists

The therapists in the Mount Zion Psychotherapy Research Project were experienced (at least 3 years of private practice) clinical psychologists and psychiatrists with a psychodynamic orientation. All had received specialized training in brief dynamic therapy. Although all of the therapists had a psychodynamic orientation, they represented several different schools of brief psychodynamic psychotherapy (e.g., the approaches of Malan, Davanloo, and Sifneos). The therapists received no information about the patients before beginning treatment; they simply knew that the patients had been screened and accepted for brief therapy. The cases reported here were studied after their therapies had been completed, and therapists were unaware of our hypotheses.

Procedure

Verbatim transcripts of Gary's and Diane's psychotherapy sessions were the source of the data for this study. The research design involved the following steps: (a) formulating the patient's therapy goals and central pathogenic beliefs, (b) locating instances of the patient testing the therapist, (c) rating the degree to which the therapist's behavior in response to the patient's tests confirmed or disconfirmed the pathogenic belief, (d) measuring the patient's behavior immediately before and after the test sequence, and (e) correlating patient changes (from pre- to posttest) with ratings of therapist behaviors.

Measures

Identifying Patient Tests

Formulation of the Pathogenic Beliefs

Clinical case formulations for these (and other) cases were developed as part of prior research on the reliability of dynamic formulations

based on control-mastery theory (Curtis et al., 1988; Curtis et al., in press; Rosenberg et al., 1986). These formulations included a description of the patient's problems and history as well as the patient's goals for therapy, the central pathogenic beliefs impeding the attainment of goals, the ways in which the patient was likely to test the therapist, and the insights that were most likely to be helpful to the patient to disconfirm pathogenic beliefs.

A brief synopsis of the cases and formulations is presented below.

Gary. A 36-year-old graduate student, Gary, sought psychotherapy because he felt inhibited, and constricted and had a sense of low self-esteem. He reported being happily married and having three children. The major precipitating factor that led him to seek therapy was critical feedback that he received in his graduate program. Several of his teachers and peers told him that he was withdrawn, too contemplative, and overly hesitant in expressing himself. He felt these criticisms were accurate and acknowledged that he had a long history of feeling inhibited, withdrawn, and self-conscious. Gary was the second of three children born to a poor, working-class family. He recalled his early childhood as a grim struggle to survive; his parents were extremely poor, unable to find work, and hopeless about the future. In school, Gary always excelled and clearly stood out among his peers. At age 14, he was offered a scholarship to a college preparatory boarding school. It was at this point in his life that his withdrawal, inhibition, and self-consciousness became prominent. Gary related his inhibition to his excelling and standing out, stating that he felt uncomfortable being regarded as "one of the best." His inhibition and depressive affect persisted throughout college and into his first job. In his job, he chronically felt like an observer rather than a participant. He felt angry and frustrated in recounting these feelings because, despite his awareness of the pattern, it had persisted.

In the formulation developed for Gary, it was inferred that he suffered from intense unconscious guilt toward his parents. He saw them as vulnerable and involved in a constant struggle to get by; he believed that they also needed him to idealize them. As a result of his guilt, he inhibited himself and made his life into a constant struggle like that of his parents. He was living out a powerful unconscious identification with them as a solution to the intense guilt he felt about being better off, happier, and more accomplished than they. He also was dealing with his guilt by complying with what he felt their expectations were, particularly their need to be revered, rather than by seeing them clearly or being critical of them. Thus, many of Gary's pathogenic beliefs were of the form that his strengths, abilities, and successes could intimidate, hurt, or overwhelm others. He inhibited the expression of his talents and identified with those weaker than him to avoid hurting them.

Diane. A 34-year-old attorney, Diane, entered therapy because of problems in motivating herself to find work. She lacked self-confidence and was plagued by a fear that she would be bullied and humiliated by other attorneys in the courtroom. Diane was the younger of two children born to a chronically unhappy mother and an alcoholic father. She described her mother as a secretive, very insecure woman who tended to ignore her daughter's feelings and would, for instance, tell Diane she was happy when Diane said she was sad. Although Diane's father was more supportive and loving, he was basically passive and had little control over his wife. While she was in high school, Diane's family experienced many problems (her brother was in prison, and her parents had serious marital problems). She described herself as depressed and withdrawn during this time. She went away to college but dropped out during her second year to get married. She wanted a divorce after the first year of marriage but remained married for 5 years because she did not want to admit she made a mistake. In the marriage, she again became withdrawn and depressed. After her divorce, she returned to college, developed interest in the law, and worked for 2 years in a law-related field before starting law school. All

during this time, she worried about losing control and humiliating herself—feelings that plagued her as she started her search for a job.

As noted earlier, Diane was seen to share many conflicts with Gary—most notably an inhibition of abilities to avoid intimidating or hurting others. In the formulation developed for Diane, it was inferred that her problems stemmed from an identification with her weak, downtrodden parents, in particular her mother. This identification developed as a consequence of her unconscious belief that if she were assertive and successful it would intimidate and hurt her parents by highlighting their inadequacies and lack of accomplishments. As a result, she was inclined to act like a victim to avoid the feeling of being the victimizer. Thus, on the threshold of becoming an attorney, Diane acted paralyzed and fragile (like her mother) and undermined her career to avoid the risk of further deflating and intimidating her parents by her success. Broadly stated, Diane's pathogenic beliefs were that her wishes, needs, and concerns could hurt others. She unconsciously believed that if she was strong or assertive, she risked hurting others by intimidating or running over them. Thus, she felt she must tightly control and inhibit her behaviors, which she did by identifying or complying with those whom she feared she would hurt by her actions.

Locating Tests

Five experienced clinicians read verbatim transcripts of each therapy session and identified all possible instances in which the patient might be testing the therapist. Each hour was read by each judge independently, thereby minimizing any systematic bias of one judge and generating the maximum number of testing sequences. In the Gary case, 45 tests were identified; 69 tests were identified in the Diane case.

Assessing the Therapist's Response to Tests

Patient Test Scale

Each patient test was excerpted from the transcript, randomized, and then presented to a new group of four experienced clinicians. Using the case formulation as a guide, the judges rated each test for the degree to which it represented a key test; that is, an instance of the patient testing a central pathogenic belief in the therapeutic relationship. A 7-point Likert scale was used, ranging from a score of 0 (indicating that the segment was not an example of the patient testing the therapist) to a score of 6 (indicating that this was an excellent, clear-cut example of the patient testing a central pathogenic belief). Only those segments with a mean score greater than 3 were regarded as key tests and included in the data analyses. For Gary, 40 key tests were identified; for Diane, there were 65.

Therapist Scale

The same judges were also asked to rate the degree to which the therapist passed or failed the patient's test (i.e., the extent to which the therapist's behavior disconfirmed or confirmed the patient's central pathogenic beliefs). These ratings were based on the case formulation and were made on a 7-point Likert scale ranging from *strongly passing the test* (6) to *strongly failing the test* (0).

Measuring the Patient's Immediate Response

The patient's immediate therapeutic progress was assessed by rating segments of his or her verbalizations immediately before (pretest, or baseline, segment) and immediately after (posttest, or effect, segment) each testing sequence. The segments of patient speech were approximately 3 min in length. These segments were isolated from the tran-

script and presented to judges randomly so that the judges were blind as to where the segment occurred in the therapy session and to whether the segment was a pre- or posttest segment; judges were also blind to treatment outcome. Three separate groups of judges independently rated each patient on the following three rating scales.

Experiencing

The patient Experiencing Scale was designed to "capture the essential quality of a client's involvement in psychotherapy" (Klein, Mathieu-Coughlan, & Kiesler, 1986, p. 21). The 7-point Experiencing Scale measures such constructs as insight, lack of resistance, and productive free association (Kiesler, 1973; Luborsky & Spence, 1978). It has been used in studies of psychoanalysis and psychoanalytic therapy (Luborsky & Spence, 1978; Silberschatz, 1978, 1986). Derived from a client-centered framework, the scale can be applied to either tape recordings or transcripts. Three psychology graduate students underwent the standardized training and rated all pre- and postsegments.

Boldness

The Boldness Scale (Caston, Goldman, & McClure, 1986) is a 5-point Likert scale that assesses a patient's ability to confront "nontrivial material." At the low end of the scale, the patient's behavior is anxious and inhibited, and at the high end, the patient is able to plunge ahead, boldly tackling new material and confronting a variety of important issues. The scale has been used in previous studies with good reliability and validity (Caston et al., 1986; Silberschatz, 1978, 1986). Predoctoral psychology interns were trained in the use of the scale and rated all pre- and posttest segments.

Relaxation

The Relaxation Scale (Curtis, Ransohoff, Sampson, Brumer, & Bronstein, 1986) measures the patient's "psychic state of freedom, relaxation, and comfort versus that of anxiety, drivenness, and beleaguering" (p. 200). It is a 5-point Likert scale originally developed for rating whole therapy sessions but adapted for use on brief segments of speech (Silberschatz, 1978, 1986). Low scores reflect evidence of the patient being tense, stressed, rigid, anxious, or driven. High scores are given when the patient is relaxed, playful, able to experience a wide range of feelings, and able to associate freely. A group of predoctoral psychology interns were trained in the use of the scale and rated all of the segments for both cases.

Reliabilities

Interjudge reliability for all rating scales was assessed by means of intraclass correlations (Shrout & Fleiss, 1979). Two figures are reported: the estimated reliability of the average judge, $r_{(j)}$ (termed ICC 3,1 by Shrout & Fleiss), and the estimated reliability of the mean of k judges' ratings, $r_{(k)}$ or coefficient alpha (termed ICC 3,K by Shrout & Fleiss). Because all subsequent data analyses involve the use of the mean ratings, $r_{(k)}$ is the appropriate measure of reliability.

Reliability results are presented for each case separately in Table 1. The obtained interjudge reliabilities were deemed sufficient for subsequent analyses. However, the reliability for the Boldness Scale for the Gary case was only marginally acceptable at .62. The low reliability figure for the Patient Testing Scale on the Diane case is attributable to a severe constriction of range (most of the tests were given high scores, indicating they were key tests) and is consequently misleading. A percentage of agreement figure was calculated to determine the absolute level of agreement among judges. The judges agreed within one scale

Table 1
Interjudge Reliabilities

Scale	Intraclass correlations			
	Diane		Gary	
	$r_{(j)}/r_{(k)}$	No. items/ no. judges	$r_{(j)}/r_{(k)}$	No. items/ no. judges
Therapist	.46/.81	69/4	.46/.77	45/4
Patient Testing	.17/.50	69/4	.42/.75	45/4
Relaxation	.47/.73	91/3	.43/.79	124/5
Boldness	.46/.72	91/3	.29/.62	124/4
Experiencing	.62/.83	91/3	.53/.77	124/3

Note. $r_{(j)}$ is the estimated reliability of the average judge; $r_{(k)}$ is the estimated reliability of K judges' ratings (coefficient alpha).

point in 84% of the tests, indicating an adequate level of interjudge agreement.

RESULTS

To assess the relationship between the therapist's behavior (the degree to which he passed or failed tests) and immediate changes in the patient, we calculated residualized gain scores (Cohen & Cohen, 1975) for ratings on the Experiencing, Boldness, and Relaxation scales. The residualized gain scores measure the variance in the postscore not predicted by the prescore. To ensure that errors of observation were independently distributed (that is, to test serial dependency), we calculated the Durbin-Watson statistic for each of the residualized change scores (for the Experiencing, Relaxation, and Boldness scales). In all instances, the results were nonsignificant, indicating that there was no serial dependency.

Using a semipartial correlation (Cohen & Cohen, 1975), we correlated the residualized gain scores with the therapist ratings. We found positive, significant correlations ($p < .01$) with the Experiencing Scale for both cases (see Table 2); that is, when the therapist passed a key test (disconfirmed a central pathogenic belief), the patient's level of experiencing increased, and

when the therapist failed a key test, the patient's level of experiencing decreased. Similar results were obtained for the Relaxation and Boldness scales for Diane but not for Gary (see Discussion section).

In addition to assessing the immediate impact of the therapist's passing or failing tests, we assessed the cumulative impact of the therapist's behavior over time. To measure this, we averaged therapist scores and the residualized scores on the Experiencing, Boldness, and Relaxation scales by hour. The mean therapist scores for each hour were then correlated with each of the mean patient residualized scores for each hour. All of these correlations were statistically significant for Diane, and the correlations with experiencing were significant for Gary (see Table 2). Graphs of the mean-per-hour therapist score and residualized experiencing scores for the cases of Diane and Gary are shown in Figures 1 and 2, respectively.

Clinical Examples

To illustrate a testing sequence, we include two brief excerpts from the case of Diane. The first is an example of a passed test, and the second is an example of a mildly failed test. Each example includes the test segment and the immediate patient response. Therapists' remarks are set off by slash marks.

Example 1

Test Segment

(silence) . . . I was just thinking about how I would cope with all my problems . . . that of course are going to come up as soon as therapy ends and . . . you know, they've all been hiding, right, for 4 months. And they'll all spring out, but I mean I'll just do it the same way I've been doing it. It's not a catastrophe; it's not the end of the world. How shall I sort this out, whatever it is? (pause)

/The problems that come up in the last couple of weeks, you seem quite comfortable in solving on your own./

Postsegment

Yeah. And I did. Yes, when I keep working something up to stand in front of me that will be the real test, which of course I'll fail. . . like going in, observing in court, or going to the clinic last Thursday night and I . . . it seems to me I could put up another one in

Table 2
Correlations Between Ratings of Therapist Behaviors and Immediate Patient Change

Patient	Scale		
	Experiencing	Boldness	Relaxation
Diane			
Individual tests ($n = 65$)	.35**	.45**	.37**
Hourly means ($n = 15$)	.67**	.57*	.55*
Gary			
Individual tests ($n = 40$)	.40**	.13	.11
Hourly means ($n = 12$)	.62*	.18	.06

* $p < .05$. ** $p < .01$.

Diane Mean-Per-Hour Therapist Scale and Residualized EXP

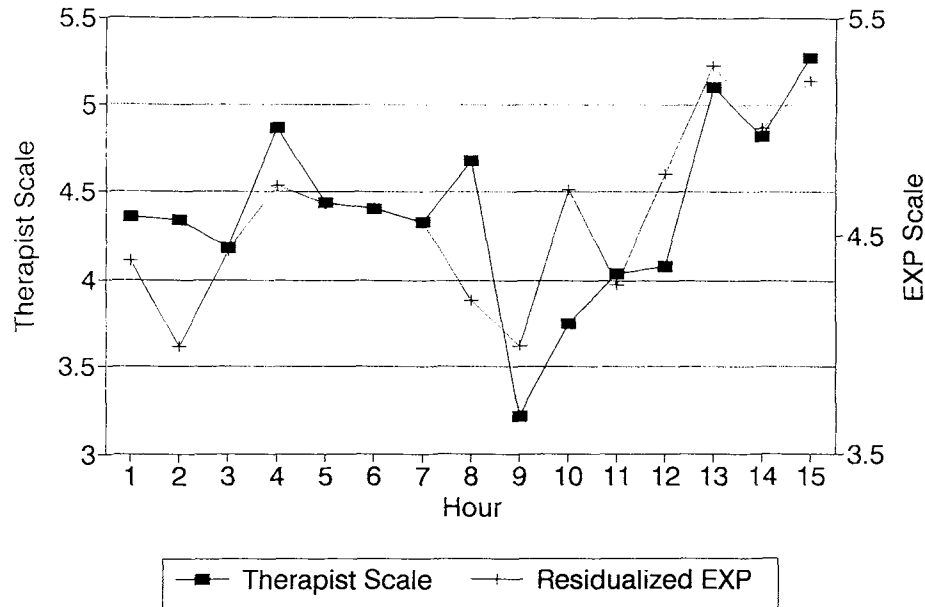


Figure 1. Experiencing Scale and Therapist Scale mean hourly ratings for Diane. (EXP = experiencing)

front of that, you know, having a trial, but I mean they're all just steps now I think, they're not road blocks. (pause) So even if I hadn't gone to the clinic on Thursday night I could still go next month for example, if I hadn't. . . I mean it would not have meant the failure of my whole life, if I'd lost my nerve or thrown up or something like that, you know it wouldn't have been the end of everything. I mean actually I had a choice Thursday, I mean, I could have gone or stayed home, or done something else. (pause, clears throat)

Comment

The plan formulation for this case (as stated earlier) indicated that Diane suffered from the pathogenic belief that her strength and success was hurtful or intimidating to others, especially her parents. In the example just presented, Diane was testing this belief by acting needy and fragile. The therapist was rated as passing this test (test passing score, 5.6 on a scale of 0 to 6) because he clearly acknowledged the patient's competence, thereby undermining her pathogenic belief that he would be bothered by her strength. After this test, the patient's levels of boldness, relaxation, and experiencing all increased relative to her baseline (pretest) ratings (residualized gain scores: Boldness Scale, 0.56; Relaxation Scale, 0.96; Experiencing Scale, 1.08).

Example 2

Test Segment

I thought I would be more, uh, frightened when this happened, or more lonely. I think we talked about that some, you know, that I would feel deserted and alone and everything. But I don't. I don't. And I thought that I would hate my mother, end up in some huge confrontation with her. And instead uhm, it just seems as though

something has adjusted definitely. But easily, you know? She, her remarks are, and her feelings . . . which she's still expressing 'cause I'm noticing some of them . . . but they just sort of are sliding off my back. I mean I feel, "Oh, really?" Or she feels differently or something, but I don't feel uhm, that I have to do anything about it, or that even that I will (clears throat). I feel embarrassed to think this is uhm, a mother problem, but it is.

/Why would you feel embarrassed about that? It just seems so uhm, so childish, and then she's just had such a great influence on me, or I've let her, or whatever. But I mean, just really really large influence. That I'm embarrassed about that. (sigh)

/How do you see that influence now? What's your kind of current kind of thinking and understanding about it?/

About?

/About the large influence she DID have on you./

Postsegment

(SIGH, pause) I'm still not sure why that happened, or why it was so. (sigh) But it seems to me now that uhm, the reason it was . . . one reason it was so large was because it was crippling for me, the way I was using it or the way I was (pause) being influenced. I would, I was always looking towards her reaction for something. And she's not consistent in her reactions in the first place, but even that aside. So there was just this constant me looking towards her reaction, and so then I would never know what I was doing. It was like a little girl I read about in a book who never learned the multiplication tables. She used to just look at her teacher's face and try to guess. So she never really learned anything; she didn't even have a place to start, with 2-by-2; you know, it was all looking at somebody else. And then you lose confidence and you . . . I lose confidence, and then I would screw something up, and I mean it just . . . just kept on and on. And I look back and I think that I was never THERE, really, as a personality. I was just there as

Gary Mean-Per-Hour Therapist Scale and Residualized EXP

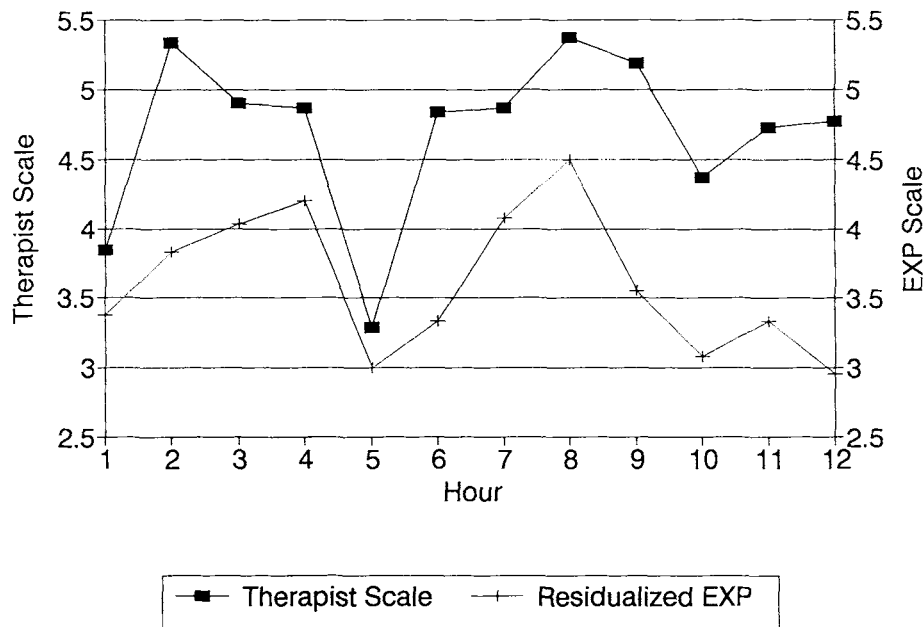


Figure 2. Experiencing Scale and Therapist Scale mean hourly ratings for Gary. (EXP = experiencing.)

a uhm, in a way as a sort of a, I don't know what, a watcher, an observer, or something like that.

Comment

In this example, the patient again tested the therapist to see if he would recognize her accomplishment (that she was less vulnerable to her mother's remarks). The therapist focused on the influence Diane's mother did have and overlooked the patient's feelings of accomplishment; hence his intervention was rated as mildly failing her test (test passing score, 2.9). As predicted, the patient was less productive after this test; there was an immediate drop on two of the three process measures (residualized gain scores: Boldness Scale, -1.35 ; Relaxation Scale, -0.35 ; Experiencing Scale, 0.12).

In these two clinical examples of testing, it was important for the therapist to acknowledge Diane's competence and strength. When he did so in Example 1, she showed signs of immediate improvement. However, it must be emphasized that this focus is specific to this case. For another patient with different issues, the therapist's focusing on the patient's strengths might be experienced as an unwillingness or inability to tolerate the patient's problems and thus fail important tests.

DISCUSSION

In this research, we address a fundamental problem in the psychotherapy research and clinical literature: how the therapist's behavior in the therapy session influences the patient's therapeutic progress. We have described a method for identify-

ing critical incidents in a therapy, for reliably judging the appropriateness of the therapist's response(s) to these incidents, and for reliably measuring the patient's reaction to the therapist's intervention. We have illustrated in this article and in others (e.g., Silberschatz, 1986; Silberschatz, Curtis, Sampson, & Weiss, 1991; Silberschatz, Fretter, & Curtis, 1986; Silberschatz, Sampson, & Weiss, 1986) that a patient does have immediate and predictable responses to a therapist's interventions and that these responses are determined to a significant degree by the "appropriateness" or "suitability" of the therapist's behavior to the patient's particular problems and needs. Broadly stated, when a therapist's interventions are in accord with the patient's goals for therapy and disconfirm the pathogenic beliefs that have inhibited the patient's progress toward achieving these goals, the patient will show signs of immediate improvement that in turn appear to contribute to outcome (Norville, 1990; Silberschatz, Fretter, & Curtis, 1986).

Our data support the widely held assumption that the therapeutic relationship is an important factor in psychotherapy, and they provide a new way of conceptualizing how the relationship may benefit the patient. In previous research, the therapeutic relationship has typically been regarded as a "non-specific factor" (Parloff et al., 1978, p. 251). Sampson (1992a, 1992b) has described how the testing concept adds a high degree of specificity to understanding the nature of the therapeutic relationship. He has stated that the therapist's attitude and behavior toward the patient—and the therapeutic relationship in general—are linked to therapeutic progress only to the extent that they disconfirm a pathogenic belief of the patient: "There is a lawful relationship in treatment between *specific experiences* which

tend to disconfirm a pathogenic belief and *specific changes* in the patient which follow from the disconfirmation of the belief" (Sampson, 1992b, p. 522).

Our results indicate that when the therapist passed the patient's key tests by responding in a manner that disconfirmed the pathogenic belief, the patient displayed immediate improvement. In the case of Diane, this improvement was reflected on all three measures; for Gary, shifts were only apparent on the Experiencing Scale. We believe that these different results reflect and are a result of the different styles of working in therapy that these patients displayed. As noted earlier, the pathogenic beliefs of Gary and Diane were broadly similar, but they displayed quite different styles of testing. Diane engaged primarily in transference testing; that is, she behaved boldly and insightfully in the therapy to test whether the therapist would be bothered by this behavior as (in her perception) her parents were. On the other hand, Gary engaged in more passive-introactive testing, in which he acted monotonously defeated, pessimistic, and un insightful, much like his parents. Throughout the course of his treatment, he had few positive comments about the therapist or the therapy, unlike Diane, who was an enthusiastic and active participant. In retrospect, we believe that the Boldness and Relaxation Scales were inappropriate for studying Gary because with his particular testing style, he would be unlikely to show much change on them.

Our experiences with the Gary case highlight issues concerning case specificity in the selection and design of patient process measures. To date, most process measures have been generic: They are applied and interpreted in the same way with all cases, usually without consideration of their clinical relevance to a given case. Determining whether a generic process measure is appropriate to a particular case requires a thorough formulation of the case that identifies what sorts of changes would be appropriate and how they are likely to be manifested. Generic measures can be thus applied in a case-specific fashion, but they may not be sufficiently sensitive to the individual needs and issues of a given case. Consequently, case-specific process measures may be required. An example of such a measure is one that we are developing called Plan Progressiveness. This measure, which is based on the formulation developed for a case, rates the degree to which a patient is discussing material that is in accord with that individual's plan ("plan progressive") versus material that may be counter to or irrelevant to the plan. As such, it is sensitive to the individual issues of a patient and how progress on these issues is likely to be manifested (see Messer, 1991; Persons, 1991, for further discussions of these issues).

This research was designed to test specific hypotheses generated by a particular theory of therapy, but the methods we have developed readily lend themselves to the study of other therapeutic modalities and to testing competing hypotheses concerning the process of psychotherapy. For example, researchers with different views of a given case can identify critical incidents and then apply their respective formulations of the case to the measurement of the therapist's interventions (see Silberschatz, 1978; Silberschatz, Sampson, & Weiss, 1986, for examples of such a study). Recently (Curtis & Silberschatz, 1990; also see Messer, 1991), investigators from two psychotherapy research groups independently developed formulations of the

Diane case. The two groups adhered to different psychodynamic theories of therapy, and the formulations they developed reflected these differences. They used each of the formulations to rate the therapist's test passing, and they used the same methods and data reported here. This strategy highlighted the relative power of each formulation to predict the patient's responses to therapist interventions. The use of this methodology across a series of cases will help in identifying commonalities and differences between theories of therapy and in testing their clinical significance.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev). Washington, DC: Author.
- Caston, J., Goldman, R. K., & McClure, M. M. (1986). The immediate effects of psychoanalytic interventions. In J. Weiss, H. Sampson, & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 277-298). New York: Guilford Press.
- Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Wiley.
- Collins, W. B., & Messer, S. B. (1988, June). *Development of case formulations employing the Mount Zion Plan Formulation Method*. Paper presented at the 19th Annual Society for Psychotherapy Research Meeting, Santa Fe, NM.
- Collins, W. B., & Messer, S. B. (1991). Extending the plan formulation method to an object relations perspective: Reliability, stability, and adaptability. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3, 75-81.
- Crits-Christoph, P., Cooper, A., & Luborsky, L. (1988). The accuracy of therapists' interpretations and the outcome of dynamic psychotherapy. *Journal of Consulting and Clinical Psychology*, 56, 490-495.
- Curtis, J. T., Ransohoff, P., Sampson, F., Brumer, S., & Bronstein, A. (1986). Expressing warded-off contents in behavior. In J. Weiss, H. Sampson, & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 187-205). New York: Guilford Press.
- Curtis, J. T., & Silberschatz, G. (1986). Clinical implications of research on brief dynamic psychotherapy: I. Formulating the patient's problems and goals. *Psychoanalytic Psychology*, 3, 13-25.
- Curtis, J. T., & Silberschatz, G. (1990, June). An empirical test of psychotherapy process predictions derived from alternative psychodynamic formulations. In J. T. Curtis (Moderator), *Applications of a model for studying and comparing theories of psychotherapy*. Panel conducted at the 21st Annual International Meeting of the Society for Psychotherapy Research, Wintergreen, VA.
- Curtis, J. T., Silberschatz, G., Sampson, H., & Weiss, J. (in press). The Plan Formulation Method. *Psychotherapy Research*.
- Curtis, J. T., Silberschatz, G., Sampson, H., Weiss, J., & Rosenberg, S. E. (1988). Developing reliable psychodynamic case formulations: An illustration of the plan diagnosis method. *Psychotherapy*, 25, 256-265.
- Fiske, D. W. (1977). Methodological issues in research on the psychotherapist. In A. S. Gurman & A. M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 23-46). New York: Pergamon Press.
- Greenberg, L. S., & Pinsof, W. M. (1986). Process research: Current trends and future perspectives. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 3-20). New York: Guilford Press.
- Horowitz, L. M., Sampson, H., Siegelman, E. Y., Wolfson, A. W., &

- Weiss, J. (1975). On the identification of warded-off mental contents. *Journal of Abnormal Psychology, 84*, 545-558.
- Jones, E. E., Cummings, J. D., & Horowitz, M. J. (1988). Another look at the nonspecific hypothesis of therapeutic effectiveness. *Journal of Consulting and Clinical Psychology, 56*, 48-55.
- Kiesler, D. J. (1973). *The process of psychotherapy: Empirical foundations and systems of analysis*. Chicago: Aldine.
- Klein, M. H., Mathieu-Coughlan, P., & Kiesler, D. J. (1986). The experiencing scales. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 21-72). New York: Guilford Press.
- Lambert, M. J., & Bergin, A. E. (1983). Therapist characteristics and their contribution to psychotherapy outcome. In C. E. Walker (Ed.), *The handbook of clinical psychology* (pp. 205-241). Homewood, IL: Dow-Jones Irwin.
- Lambert, M. J., Shapiro, D. A., & Bergin, A. E. (1986). The effectiveness of psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 157-211). New York: Wiley.
- Luborsky, L., & Spence, D. (1978). Quantitative research on psychoanalytic therapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 331-368). New York: Wiley.
- Messer, S. B. (1991). The case formulation approach: Issues of reliability and validity. *American Psychologist, 46*, 1348-1350.
- Norville, R. (1990). Plan compatibility of interpretations and brief psychotherapy outcome (Doctoral dissertation, Pacific Graduate School of Psychology, 1989). *Dissertation Abstracts International, 50*, 12B.
- Parloff, M. B., Waskow, I. E., & Wolfe, B. E. (1978). Research on therapist variables in relation to process and outcome. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (2nd ed., pp. 233-282). New York: Wiley.
- Perry, J. C., Luborsky, L., Silberschatz, G., & Popp, C. (1989). An examination of three methods of psychodynamic formulation based on the same videotaped interview. *Psychiatry, 52*, 302-322.
- Persons, J. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist, 46*, 99-106.
- Rice, L. N., & Greenberg, L. S. (1984). The new research paradigm. In L. N. Rice & L. S. Greenberg (Eds.), *Patterns of change* (pp. 7-26). New York: Guilford Press.
- Rosenberg, S. E., Silberschatz, G., Curtis, J. T., Sampson, H., & Weiss, J. (1986). A method for establishing reliability of statements from psychodynamic case formulations. *American Journal of Psychiatry, 143*, 1454-1456.
- Sampson, H. (1992a, June). An overview of some implications of our findings. In G. Silberschatz (Moderator), *The San Francisco Psychotherapy Research Group: An overview of a decade of work (Part II)*. Panel conducted at the 23rd Annual International Meeting of the Society for Psychotherapy Research, Berkeley, CA.
- Sampson, H. (1992b). The role of "real" experience in psychopathology and treatment. *Psychoanalytic Dialogues: A Journal of Relational Perspectives, 2*, 509-528.
- Schaffer, N. D. (1982). Multidimensional measures of therapist behavior as predictors of outcome. *Psychological Bulletin, 92*, 670-681.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.
- Silberschatz, G. (1978). Effects of the analyst's neutrality on the patient's feelings and behavior in the psychoanalytic situation. *Dissertation Abstracts International, 39*, 3007B. (University Microfilms No. 78-24, 277)
- Silberschatz, G. (1986). Testing pathogenic beliefs. In J. Weiss, H. Sampson, & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 256-266). New York: Guilford Press.
- Silberschatz, G., & Curtis, J. T. (1986). Clinical implications of research on brief dynamic psychotherapy: II. How the therapist helps or hinders therapeutic progress. *Psychoanalytic Psychology, 3*, 27-37.
- Silberschatz, G., Curtis, J. T., & Nathans, S. (1989). Using the patient's plan to assess progress in psychotherapy. *Psychotherapy, 26*, 40-46.
- Silberschatz, G., Curtis, J. T., Sampson, H., & Weiss, J. (1991). Mount Zion Hospital and Medical Center: Research on the process of change in psychotherapy. In L. E. Beutler & M. Crago (Eds.), *Psychotherapy research: An international review of programmatic studies* (pp. 56-64). Washington, DC: American Psychological Association.
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology, 54*, 646-652.
- Silberschatz, G., Sampson, H., & Weiss, J. (1986). Testing pathogenic beliefs versus seeking transference gratifications. In J. Weiss, H. Sampson, & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 267-276). New York: Guilford Press.
- Stiles, W. B., Shapiro, D. A., & Elliott, R. (1986). "Are all psychotherapies equivalent?" *American Psychologist, 41*, 165-180.
- Strupp, H. H. (1986). Psychotherapy: Research, practice, and public policy (How to avoid dead ends). *American Psychologist, 41*, 120-130.
- Weiss, J. (1986). I. Theory and clinical observations. In J. Weiss, H. Sampson, & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 3-138). New York: Guilford Press.

Received February 26, 1992

Revision received July 10, 1992

Accepted October 16, 1992 ■