



One step further in assessing how interpretations influence the process of psychotherapy

F. Caspar , J. Pessier , J. Stuart , J.D. Safran , L.W. Samstag & M. Guirguis


To cite this article: F. Caspar , J. Pessier , J. Stuart , J.D. Safran , L.W. Samstag & M. Guirguis (2000) One step further in assessing how interpretations influence the process of psychotherapy, *Psychotherapy Research*, 10:3, 309-320, DOI: [10.1080/713663734](https://doi.org/10.1080/713663734)

To link to this article: <https://doi.org/10.1080/713663734>



Published online: 23 Apr 2010.




[Submit your article to this journal](#) 



Article views: 100



[View related articles](#) 



Citing articles: 8 [View citing articles](#) 

ONE STEP FURTHER IN ASSESSING HOW INTERPRETATIONS INFLUENCE THE PROCESS OF PSYCHOTHERAPY

Franz Caspar
University of Freiburg, Germany
Julian Pessier
Jennifer Stuart
Jeremy D. Safran
New School for Social Research
Lisa Wallner Samstag
Manal Guirguis
Beth Israel Medical Center

The question of how interpretations take effect is among the most important in psychotherapy research. In a landmark study, Silberschatz, Fretter, and Curtis (1986) developed a method to assess the compatibility between an interpretation and a patient's psychology (the patient's unconscious plan), and demonstrated that this factor—plan compatibility—may contribute substantially to an interpretation's effect. However, they did not consider the following possibility: An interpretation might be plan compatible in terms of content, but incompatible in terms of process (e.g., an interpretation accurately identifying a patient's fear of acting independently may be delivered in a controlling way), or vice-versa. In a reanalysis of data provided by Silberschatz et al. (1986), we demonstrate that raters can reliably distinguish content from process in evaluating plan compatibility. We demonstrate, further, that a distinction between content and process plan compatibility may increase the amount of explained variance in Patient Experiencing scores. Of particular interest is an analysis of interpretations in which the aspects of plan compatibility (content or process) differed, which showed that the aspect of plan compatibility most highly correlated with Experiencing varied across the cases. This finding suggests that attention to the distinct contributions of content and process to plan compatibility may help further our understanding of how individual patients listen to and utilize therapist interventions.

One advantage of a systematic case formulation method is to permit evaluation of therapist interventions with reference to specific, articulated standards. To the extent that a patient's wishes, fears, and therapeutic goals can be detected and speci-

The authors thank Elisabeth Schneeberger for her help with the statistics.

Correspondence concerning this article should be addressed to Franz Caspar, Ph.D., Institute of Psychology, University of Freiburg, Belfortstr. 18, 79098 Freiburg i.Br., Germany.

fied, the suitability of a therapist's efforts to help can be evaluated. This principle guides work reported by Silberschatz, Fretter, and Curtis (1986), in which a formulation method developed by the Mount Zion Psychotherapy Research Group is used to identify *plans*, composed of both conscious and unconscious elements, which the authors believe guide patients' efforts to solve their problems through treatment. The Plan Diagnosis method was used to gauge the *plan compatibility* of therapists' interventions, i.e., how well therapists' interventions fit their patients' particular needs. The authors demonstrated significant positive correlations between plan compatibility and patient productivity, as indicated by Patient Experiencing Scale ratings (Klein, Mathieu-Coughlan, & Kiesler, 1986) in three cases of brief psychodynamic psychotherapy. Results of their study also suggested that plan compatibility may be more useful than category of interpretation (transference versus nontransference) in identifying efficacious therapist interventions. Despite limitations associated with small sample size and reliance upon a single dependent measure, the Silberschatz et al. (1986) study makes an important stride in the direction of taking quality of intervention and contextual appropriateness into account in psychotherapy process research. Many subsequent citations of the paper (e.g., Crits-Christoph, Barber, Baranackie, & Cooper, 1993; Crits-Christoph, Cooper, & Luborsky, 1988; Piper, McCallum, Azim, & Joyce, 1993; Spiegel & Hill, 1989) attest that it has become a seminal work in the field of research on process and outcome in psychodynamic therapy.

Its seminal status in the field makes the Silberschatz et al. (1986) study worthy of continued scrutiny. We propose that the definition of plan compatibility remains ambiguous on one important count: Is plan compatibility of therapists' interventions evaluated with respect to content or process? In a more recent study, using different patients than in the original study, Silberschatz and Curtis (1993) assessed the plan compatibility of therapist behavior in response to patient-initiated tests of pathogenic beliefs. This later study defined plan compatibility in a manner similar to what we define as the process component of plan compatibility. However, this attempt to study the effects of process plan compatibility independently of global plan compatibility leaves our original question intact: What are the relative contributions of content and process to global plan compatibility? Silberschatz et al. acknowledge (personal communication, 1996) that, to date, the potentially separate contributions of content and process to plan compatibility have not been teased apart.

One important goal of the current study was to establish whether content and process can be reliably distinguished in evaluating plan compatibility. Additional goals, contingent upon successful achievement of the first, were to establish whether one or the other aspect of plan compatibility has greater bearing upon (a) plan compatibility ratings reported in the original Silberschatz et al. (1986) study, and (b) patient progress, as reflected in Patient Experiencing Scale (Klein et al., 1986) ratings in the original study. If we can establish differential contributions of content and process to plan compatibility and to patient progress, we might be in a position to specify further how interpretations influence the therapy process. Our study also may be viewed as an indirect test of the transferability of the plan formulation method. Any meaningful covariation between our content and process plan compatibility ratings and the original plan compatibility ratings would suggest that our judges were able to understand and meaningfully apply the plan formulations employed in the original study.

In distinguishing between content and process, we work within an interpersonal perspective articulated by (among others) Sullivan (1953) and Kiesler (1973). Following convention in this tradition, we conceive of content as the explicit meaning

of a speaker's words. Thus, the content aspect of plan compatibility concerns how a therapist's words refer explicitly to a patient's plan. For example, for a patient working to overcome guilt feelings related to becoming independent of her mother, the following interpretation would be plan incompatible in content: "No wonder your mother is disappointed when you show so little interest in her." By process, we mean how an interpretation takes its effect(s) as an act in the context of an interpersonal relationship. The process of an interpretation may or may not correspond with its content. For example, for the same patient described above, the following interpretation might have divergent values with respect to content and process plan compatibility: "See how following my suggestions helps you to become independent of your mother, without feeling guilty?" Here, the therapist urges the patient toward guiltless independence from her mother (plan compatible content), but in the act of urging, promotes dependence upon the therapist (plan incompatible process). For the sake of clarity, we have chosen a caricaturishly stark example of divergence between content and process; however, it should serve to illustrate the distinction we have in mind. It is important to note that like judges in the Silberschatz et al. (1986) study, our judges worked from transcripts, and thus evaluated both content and process dimensions of therapists' interpretations on the basis of their words alone. Judges working with audio- or videotaped material undoubtedly could bring additional, nonverbal material to bear in distinguishing between content and process.

METHOD

SILBERSCHATZ, FRETTER, AND CURTIS, 1986

We are indebted to Silberschatz, Fretter, and Curtis for permission to reanalyze their data. Readers are referred to their 1986 paper for full description of their data collection and plan compatibility assessment procedures. Here, we remind readers only of details relevant to our reanalysis of the data.

Silberschatz et al. (1986) selected three brief dynamic therapy cases for study, on the basis of patients', therapists', and independent evaluators' ratings on standard psychotherapy outcome measures: Case 1 showed excellent outcome; Case 2, moderately good outcome; and Case 3, poor outcome. All three patients were diagnosed with "chronic, neurotic depression, or dysthymic disorder" (p. 647), and were randomly assigned to experienced therapists trained in brief dynamic therapy for a 16-week course of treatment. All sessions of each patient's therapy were recorded. Four graduate student judges read complete verbatim transcripts of each session and categorized all therapist interventions as either interpretations or noninterpretations (N). Interpretations were further categorized as concerning the transference (T); the patient's relationship with a parent or sibling (P), or with another identifiable person (O); or as undirected (U). For each case, interpretations in the transference category and in the most frequently occurring nontransference category were retained for further study.

For each case, a team of five experienced clinicians prepared a plan formulation, specifying "(a) the patient's *goals* for therapy, (b) the inner *obstacles* (pathogenic beliefs) preventing the attainment of goals, (c) the means by which the patient would *test* the therapist to disconfirm his/her pathogenic beliefs, and (d) the *insights* that would be helpful to the patient" (Silberschatz et al., 1986, pp. 647-648). For each case, a separate team of four to six experienced clinicians employed the Plan

Compatibility of Interventions Scale (PCIS; Caston, 1986) to rate each therapist intervention identified as an interpretation on an ordinal scale ranging from -3 for strongly antiplan interpretations to $+3$ for strongly proplan interpretations. Interpretations that judges could not rate with adequate agreement (within one standard deviation of each others' ratings) on the PCIS were excluded from further analysis, on the assumption that widely discrepant ratings reflected their inherent ambiguity.

The Patient Experiencing Scale (EXP; Klein et al., 1986) was used as an in-session outcome criterion. EXP ratings were made by six graduate and undergraduate students in psychology. Residualized gain scores (Cohen & Cohen, 1975) were used to compare three-minute segments of patient speech immediately preceding and immediately following interpretations. Differences in residualized gain scores were used to assess the differential effects of interpretations varying in category (transference versus nontransference) and in degree of plan compatibility.

REANALYSIS IN THE CURRENT STUDY

For the current study, we obtained the interpretations culled from each case, and the three corresponding plan formulations, from Silberschatz and colleagues (1986). The modified rating procedure, in which each interpretation is rated separately on the PCIS for content and process plan compatibility (CPC and PPC, respectively), was developed with separate case material, also obtained from the Mount Zion group. In an effort to keep content and process ratings as independent as possible, we tried initially to have the two aspects of plan compatibility rated by two separate sets of independent judges. However, this approach proved unfeasible. It seems one needs to rate both aspects of plan compatibility in order to make a clear distinction between content and process. If resources were unlimited, it might be desirable to employ two groups of judges, each of whom rates both aspects of plan compatibility, and then to use CPC ratings from one group and PPC ratings from another.

For the present study, a graduate student and an undergraduate, familiar with (though not formally trained in) the plan formulation method, served as judges. The judges worked with practice material from the case loaned for rating procedure development until they reached a level of reliability sufficient to begin the actual rating work. They then began rating Cases 1, 2, and 3. To permit ongoing evaluation of interrater reliability, judges assigned ratings independently, prior to discussion, throughout the study. Discrepant ratings were discussed and resolved by consensus; to ensure that reliability was maintained as the work proceeded, reliability checks were made on preconsensus ratings. In keeping with Silberschatz and colleagues' (1986) decision to exclude interpretations for which agreement about PCIS ratings was poor, we excluded a small number of interpretations for which judges' preconsensus ratings of CPC and PPC were extremely discrepant. An interpretation was excluded if judges' individual ratings of either content or process differed by at least four points, or if both content and process ratings differed by at least three points. By these criteria, we excluded three interpretations from Case 1; two from Case 2; and two from Case 3. A difference between content and process values of at least one point on the seven-point PCIS scale, for at least 20% of the interpretations studied, was established in advance as a criterion to evaluate whether the distinction between CPC and PPC warranted further exploration.

Rather than rely upon the residualized gain score as sole indicator of an intervention's effect upon patient experiencing, we included the experiencing rating for

the postinterpretation segment in all analyses as well. It is easy to imagine distorting influences upon the relationship between quality of interpretation and change in experiencing, measured from pre- to postintervention. For example, low preinterpretation experiencing values may indicate that a patient is not ready to make use of a plan-compatible intervention; in this case, little change in experiencing would occur, despite a good interpretation. In contrast, high experiencing prior to an interpretation might have a ceiling effect, leaving little room for a well-timed, plan-compatible intervention to show its effect. The latter example illustrates the value of considering postinterpretation experiencing ratings in addition to residualized gain; evidence of an effective interpretation, lost in the change score, would be reflected in the simple postinterpretation score.

RESULTS

INTERRATER RELIABILITY

Interrater reliabilities for CPC and PPC were calculated using intraclass correlations. To assess agreement between the two judges, we used the coefficient referred to as ICC (2,1) (Shrout & Fleiss, 1979). Table 1 shows the reliability coefficients for the individual interpretations, calculated separately for each of the three cases: For CPC, the mean of the coefficients was .57, ranging from .53 to .61; for PPC, the mean of the coefficients was .57, ranging from .48 to .64. Correlations between CPC and PPC, for each judge, were higher than correlations between judges, within aspect of plan compatibility (CPC or PPC); i.e., more variance can be attributed to rater than to aspect of plan compatibility rated. Thus, both content and process plan compatibility ratings may be determined mainly by a judge's general view of an interpretation.

The reliability figures were lower than the values reported in Silberschatz and colleagues' original study (1986), in large part because those authors utilized several judges for each rating task, which tended to boost overall reliability. Considering the difficult task posed to our raters, the current study would have benefited from the use of multiple judges, as well as from the use of more clinically experienced judges. Although the figures were lower than we had hoped, and can be considered only marginally acceptable, we chose to proceed with the proposed analyses anyway, for two reasons: (1) The goal of the current study was not to test specific clinical hypotheses, but rather to conduct exploratory analysis to evaluate whether the use of content and process plan-compatibility ratings is a worthwhile methodological endeavor; and (2) interrater reliability was high enough to suggest that with increased use and understanding of the task, future investigators wishing to utilize the measure could take steps to ensure higher coefficients (e.g., use of additional judges, more specific guidelines for distinguishing content and process).

TABLE 1. Interrater Reliabilities for CPC and PPC Ratings

	Number of interpretations	Intraclass correlation coefficients (2,1)	
		Content (CPC)	Process (PPC)
Case 1	63	.57	.48
Case 2	73	.61	.64
Case 3	64	.53	.59

VARIATION BETWEEN CONTENT AND PROCESS RATINGS

Nonetheless, judges were able to distinguish content from process for a substantial number of interpretations. The criterion of a one-point difference between CPC and PPC in at least 20% of interpretations studied was surpassed in all three cases (Table 2; data reported are for judges' consensus ratings). Notably, the percentage of interpretations for which CPC and PPC differ by at least a point corresponded directly with goodness of outcome. Content and process evidently were easiest to tease apart in Case 1, where outcome was judged to be excellent; somewhat less discrete in Case 2, where outcome was moderately good; and most difficult to distinguish in Case 3, where outcome was poor. Additional results, reported later on, will suggest that where it is possible to distinguish content and process, it may be important to do so.

ASSESSING CONTENT AND PROCESS CONTRIBUTIONS TO GLOBAL PLAN COMPATIBILITY AND EXPERIENCING

Predicting global plan compatibility. To assess the contributions of content and process to global plan compatibility, CPC and PPC ratings were correlated with global plan compatibility (GPC), as rated in the original study. To establish whether content and process had differential effects upon patient progress, CPC and PPC were correlated with EXP ratings in segments following interpretations and (as in the earlier study) with residualized EXP ratings.

Table 3 shows means and standard deviations for the basic variables. Although the data invite further reflections, we concentrate our evaluation and interpretation—parallel to Silberschatz et al. (1986)—on differences in correlation values.

Following Silberschatz and colleagues (1986), we conducted separate analyses for individual interpretations and for hourly means within each case. Thus we were able to explore relationships of CPC and PPC with global plan compatibility and patient progress at both an immediate, moment-to-moment level and a more molar, session-to-session level. Both sets of correlations are reported in Table 4.

Correlations of CPC and PPC with GPC ratings provided by Silberschatz and colleagues (1986) generally were moderate to high, and most were significant. That this was so speaks for the exportability of the Mount Zion method. Our raters' overall agreement with raters in the original study suggested that though they lacked extensive training in the approach, they were able to understand the concept of plan diagnosis, in general terms and as articulated in the content of each case.

Across the three cases, GPC correlated to comparable degrees with CPC and PPC; thus, it was not possible to say that either content or process consistently held most sway in determining plan compatibility ratings. In Case 1, correlations between PPC and GPC tended to be slightly larger than correlations between CPC and GPC. For

**TABLE 2. Frequency With Which CPC and PPC Ratings
(Consensus) Differ by at Least One Point**

	Number of interpretations	
	Total	With CPC and PPC differing by at least one point
Case 1	63	32 (51%)
Case 2	73	29 (40%)
Case 3	64	18 (28%)

Note. CPC, content plan compatibility; PPC, process plan compatibility.

TABLE 3. Means and SDs of CPC, PPC, GPC, PEXP and REXP Consensus Ratings

	CPC		PPC		GPC		PEXP		REXP	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Case 1	.71	1.59	1.30	1.02	1.79	.67	3.31	.70	.04	.65
Case 2	1.47	.99	1.57	.85	1.67	.76	3.31	.71	.03	.70
Case 3	1.27	.81	1.27	.92	.68	1.08	3.22	.56	.01	.55

Note. CPC, content plan compatibility; PPC, process plan compatibility; GPC, global plan compatibility; PEXP, post-intervention experiencing ratings; REXP, residual experiencing ratings.

Case 2, there were small differences in the opposite direction: correlations between CPC and GPC were higher than correlations between PPC and GPC. In Case 3, correlations of GPC and CPC were sometimes higher and sometimes lower than correlations of GPC with PPC. Overall, the results from these three cases suggest that neither content nor process consistently contributes most to plan compatibility.

Predicting patient progress (EXP). Silberschatz et al. (1986) reported a substantially higher correlation of GPC with residual experiencing scores for Case 1 than for the other two cases. This pattern of correlations between plan compatibility and experiencing turned out to be remarkably consistent; it occurred for both CPC and PPC, whether one considers postinterpretation or residual experiencing scores. When correlations of CPC and PPC with EXP were considered separately, a further refinement of the pattern occurs, which remarkably coincided with outcome: Correlations of both CPC and PPC with EXP were highest in Case 1, where outcome was excellent; middling in Case 2, where outcome was moderately good; and lowest in Case 3, where outcome was poor. A comparison of correlations of CPC with EXP and PPC with EXP suggested that separate content and process ratings captured similar levels of covariance with experiencing.

TABLE 4. Correlations of CPC and PPC Ratings With Global Plan Compatibility and EXP Ratings From Silberschatz et al. (1986)

	Case	Individual interpretations			Hourly means		
		GPC	PEXP	REXP	GPC	PEXP	REXP
Content (CPC), consensus ratings	1	.47**	.52**	.43**	.77**	.90**	.86**
	2	.60**	.26*	.25*	.69*	.58*	.62*
	3	.49**	.16	.17	.55*	.04	.01
Process (PPC), consensus ratings	1	.52**	.53**	.52**	.81**	.88**	.86**
	2	.48**	.29*	.29*	.51	.49	.54*
	3	.44**	.13	.16	.59	.07	.04
Global plan compatibility (GPC)	1		.55**	.55**		.79**	.80**
	2		.28*	.27*		.61*	.59*
	3		.28*	.28*		.43	.38

Note. Global plan compatibility (GPC), post-intervention experiencing ratings (PEXP), and residual experiencing ratings (REXP) are from Silberschatz et al. (1986), and were generously provided for reanalysis by the authors of that study.

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

We were interested in the effect of using different rating types—experiencing values following interpretations (PEXP) versus residual experiencing values (REXP; calculated by subtracting preinterpretation from postinterpretation values) on the correlations observed between the types of plan compatibility and levels of experiencing. Correlations of plan compatibility with PEXP and REXP were generally similar, though instances in which they did diverge attested to the potential importance of using both. The differences in correlation between the plan compatibility variables and PEXP vs. REXP in Case 1 were hard to interpret but should draw attention to the difference between the two experiencing variables with more data in further studies.

INTERPRETATIONS WITH DIVERGENT CONTENT AND PROCESS RATINGS

The analyses reported above include data for interpretations in which content and process ratings did not differ. We next conducted the same analysis but with a less dilute sample, consisting only of interpretations for which content and process ratings differed by at least one point. These analyses brought the differential contributions of content and process more clearly into view. Because this further selection resulted in unequal frequencies among sessions, we conducted correlations only for individual interpretations (not for hourly means); results are given in Table 5.

For Case 1, GPC ratings correlated more highly with process (PPC with GPC, $r = .39$, $p < .05$) than with content (CPC with GPC, $r = -.06$, ns). However, PEXP ratings for Case 1 were more highly correlated with content (CPC with PEXP, $r = .41$, $p < .05$) than with process (PPC with PEXP, $r = .25$, ns); correlations of content and process with REXP did not differ appreciably (CPC with REXP, $r = .22$, ns; PPC with REXP, $r = .25$, ns). Furthermore, correlations between GPC and EXP were very low in this case (GPC with PEXP, $r = .07$, ns; GPC with REXP, $r = .09$, ns). Thus in Case 1, for interpretations in which CPC and PPC differ, GPC ratings may miss most of the

TABLE 5. Correlations of Consensus CPC and PPC Ratings With Global Plan Compatibility and EXP Ratings From Silberschatz et al. (1986): Selected Sample

	Case	Individual interpretations		
		GPC	PEXP	REXP
Content (CPC), consensus ratings	1	-.06	.41*	.22
	2	.55*	.25	.25
	3	.68**	.19	.11
Process (PPC), consensus ratings	1	.39*	.25	.25
	2	.38*	.40*	.41*
	3	.67**	.02	.03
Global plan compatibility (GPC)	1		.07	.09
	2		.25	.22
	3		.19	.13

Note. Global plan compatibility (GPC), post-intervention experiencing ratings (PEXP), and residual experiencing ratings (REXP) are from Silberschatz et al. (1986), and were generously provided for re-analysis by the authors of that study.

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

therapeutic gain. In this case, it is the content of the interpretation that appears to carry more of the therapeutic effect.

For Case 2, GPC ratings correlated more highly with content (CPC with GPC, $r = .55, p < .05$) than with process (PPC with GPC, $r = .38, p < .05$). However, EXP in the selected segments seemed to be determined more by process (PPC with PEXP, $r = .40, p < .05$; PPC with REXP, $r = .41, p < .05$) than by content (CPC with PEXP, $r = .25, ns$; CPC with REXP, $r = .25, ns$). Although correlations between GPC and EXP were not as low as for Case 1 (here, GPC with PEXP, $r = .25, ns$; GPC with REXP, $r = .22, ns$), they were considerably lower than correlations between PPC ratings and EXP ratings. As in Case 1, then, GPC ratings seem to miss much of the therapeutic gain. Here, however, the situation observed in Case 1 is reversed: For the selected interpretations in Case 2, therapeutic gain is determined mostly by process.

For Cases 1 and 2, correlations of content and process with GPC ratings tended to decrease with use of the selected sample. In Case 3, the use of selected interpretations increased correlations of both content and process with GPC (CPC with GPC, $r = .68, p < .01$, as compared with $r = .49, p < .05$ in the earlier analysis; PPC with GPC, $r = .67, p < .01$, as compared with $r = .44, p < .05$ in the earlier analysis), and the contributions of content and process to GPC were nearly equal. Despite the increased correlations with plan compatibility, CPC and PPC ratings continued to show little relationship with experiencing ratings in this case.

Of all reported differences in correlations, the higher correlation between CPC as well as PPC and EXP for Case 1 compared to the other cases may partly be attributed to the higher variation of the two former variables for Case 1. The higher correlation between CPC and PEXP (not with REXP) in Case 1 may partly be attributed to the higher variance of CPC. If one sees the correlations in a pattern, a consideration of possible variance effects would not lead to a significant change of the interpretations.

DISCUSSION

The implications of our main findings regarding content and process contributions to the plan compatibility of therapists' interventions can be summarized as follows:

1. Differential contributions of content and process to plan compatibility could be distinguished in a substantial portion of therapist interventions (about 40% overall). This result was obtained using transcripts of therapy sessions; use of audio- or videotape, in conjunction with transcripts, might permit such distinctions in a higher proportion of instances. Building a content-process distinction into the development of case formulations also might enhance judges' ability to distinguish these aspects of plan compatibility in therapists' individual interpretations. Our judges' ability to understand and make use of plan formulations provided by the Mount Zion group, without formal training in the Mount Zion approach and despite the lack of explicit distinction between content and process in the formulations, argues for the exportability of the plan diagnosis method.
2. Content and process ratings of plan compatibility make very similar contributions to the Mount Zion group's assessment of plan compatibility in the original Silberschatz et al. (1986) study. Some small differences in the magnitudes of correlations suggest that process may have contributed more to GPC in

Case 1, and content to GPC in Case 2, but neither content or process consistently contributed more to overall plan compatibility.

3. Differential contributions of content and process to patient progress, as reflected in patient experiencing, become apparent when one considers only those interpretations for which CPC and PPC ratings vary by at least one point. For this subsample, content and process contributions to experiencing vary from case to case. EXP is most highly correlated with content in Case 1, and with process in Case 2; in Case 3, no relationship was observed between either content or process with EXP. These findings suggest that a distinction between content and process may improve predictions of the therapeutic impact of interpretations.

Given the small size of our sample, the results reported above cannot be assumed to generalize, and must be regarded as merely suggestive. What the overall similarity in the relationships of CPC and PPC ratings with GPC and EXP suggest is that neither aspect of plan compatibility is consistently more important than the other. However, the analysis of interpretations that differ in CPC and PPC suggest that the distinction between content and process may offer the potential to sharpen our understanding of how therapists' interpretations take effect—or fail to take effect—in the context of a particular dyad.

Two incidental findings, both concerning the ordering of the three cases in terms of outcome, are of considerable interest. Content and process could be distinguished most frequently in Case 1 (excellent outcome); somewhat less frequently in Case 2 (moderately good outcome); and less frequently in Case 3 (poor outcome). This finding is somewhat counterintuitive. One might expect to find good outcome in association with high congruity between the content and process dimensions of therapists' interventions; bad process might be expected to undermine good content, and vice versa. Given the small size of the study sample, our findings may not be reliable; thus, attempts to explain them may be prematurely speculative. However, it might be prudent for future studies distinguishing between CPC and PPC to take into account the following possibility: Perhaps some mediating factor, such as the clarity or the specificity of the therapist's speech, is responsible both for the apparent effectiveness of his/her interventions and for judges' relative ease in distinguishing content from process in cases with good outcome.

Correlations of CPC and PPC with EXP also parallel more positive outcome, as follows: They are highest in Case 1 (excellent outcome), middling in Case 2 (moderately good outcome), and lowest in Case 3 (poor outcome). This finding suggests that the degree of relationship between what a therapist does and how a patient responds may function as a predictor of outcome. It would be interesting to explore this possibility with additional process measures and in additional samples.

Research involving a clinical case formulation method such as the Mount Zion approach requires extensive resources for each case studied; thus, sample size is necessarily restricted. However, the results found in the two cases for which outcome was at least moderately good raise the interesting possibility of distinctive mechanisms of action, one based on process and one on content—each effective, but by different means. In this regard, our study argues for the importance of further work, with an extended sample. The inclusion of additional process measures undoubtedly would help to improve our understanding of how content and process dimensions of plan compatibility contribute to the effectiveness of an intervention, in a particular dyadic context. For example, the Therapist Experiencing Scales (Klein

et al., 1986), with Referent and Manner subscales, hold promise as additional means to assess the therapist's performance. Furthermore, it should be noted that patient experiencing is only one of many possible indicators of an intervention's effect; the inclusion of additional patient process variables undoubtedly would contribute to our understanding as well.

In conclusion, a distinction between content and process aspects of plan compatibility has the potential to provide a more detailed picture than can GPC ratings of how therapists' interpretations take (or fail to take) their effects. Concomitant use of process measures additional to GPC and EXP would help to illuminate further the complex patterns discovered in this small sample, and, of course, a larger sample would be required to support any confident generalization. For now, we offer our findings as a demonstration that it may prove profitable to distinguish between content and process in future development of the Mount Zion plan formulation method, and indeed in case formulation methods generally.

REFERENCES

- Caston, J. (1986). The reliability of the diagnosis of the patient's unconscious plan. In J. Weiss, H. Sampson, & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observations & empirical research* (pp. 241–255). New York: Guilford Press.
- Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Wiley.
- Crits-Christoph, P., Barber, J. P., Baranackie, K., & Cooper, A. (1993). Assessing the therapist's interpretations. In N. E. Miller, L. Luborsky, J. P. Barber & J. Docherty (Eds.), *Psychodynamic Treatment Research* (pp. 361–386). New York: Basic Books.
- 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* (4), 490–495.
- 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. Pincus (Eds.), *The psychotherapeutic process: A research handbook* (pp. 21–71). New York: Guilford.
- Piper, W. E., McCallum, M., Azim, H. F. A., & Joyce, A. S. (1993). Understanding the relationship between transference interpretation and outcome in the context of other variables. *American Journal of Psychotherapy, 47* (4), 479–493.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86* (2), 420–428.
- Silberschatz, G., & Curtis, J. T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology, 61* (3), 403–411.
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *Journal of Consulting and Clinical Psychology, 54*, 646–652.
- Spiegel, S. B. & Hill, C. E. (1989). Guidelines for research on therapist interpretation: Toward greater methodological rigor and relevance to practice. *Journal of Counseling Psychology, 36* (1), 121–129.
- Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton Press.

Zusammenfassung

Die Frage, wie sich Deutungen auswirken, gehört zu den wichtigsten in der Psychotherapieforschung. In einer richtungsweisenden Studie konnten Silberschatz, Fretter und Curtis (1986) eine Methode entwickeln, mit der die Kompatibilität zwischen einer Deutung und der Psychologie des Patienten (dem unbewußten Plan des Patienten) erfaßt werden konnte. Die Autoren zeigten, daß dieser Faktor, die Plankompatibilität, substantiell zum Effekt einer Deutung beiträgt. In dieser Studie wurde aber die folgende Möglichkeit nicht in Betracht gezogen: eine Deutung mag zwar plankompatibel sein im Hinblick auf den Inhalt, aber inkompatibel im Hinblick auf den Prozeß oder vice-versa. In einer Reanalyse der

Daten, die uns von Silberschatz u.a. zur Verfügung gestellt wurden, zeigen wir, daß Beurteiler auf reliable Weise bei der Bewertung der Plankompatibilität zwischen Inhalt und Prozeß unterscheiden können. Außerdem zeigen wir, daß eine Unterscheidung zwischen Inhalts- und Prozeßplankompatibilität die erklärte Varianz in patientenbezogenen Maßen erhöhen kann. Von besonderem Interesse ist dabei die Analyse solcher Deutungen, bei denen die Aspekte der Plankompatibilität (Inhalt versus Prozeß) unterschiedlich waren. Hier zeigte sich, daß der Aspekt der Plankompatibilität der am höchsten mit dem Experiencing korreliert war, über die Fälle variierte. Dieser Befund legt nahe, daß die Bemühung, zwischen Inhalt und Prozeß im Zusammenhang mit der Plankompatibilität zu differenzieren, nützlich ist, um zu klären, wie individuelle Patienten auf Therapeuteninterventionen hören und diese nutzen.

Résumé

Comprendre comment une interprétation devient efficace fait partie des interrogations les plus importantes dans la recherche en psychothérapie. Dans une étude pionnière, Silberschatz, Fretter et Curtis (1986) avaient développé une méthode pour évaluer la compatibilité entre une interprétation et la psychologie d'un patient (le *plan* inconscient du patient), en démontrant que ce facteur—la *compatibilité avec le plan*—peut contribuer substantiellement à l'effet d'une interprétation. Ils n'ont cependant pas tenu compte de la possibilité suivante : une interprétation peut être compatible avec le plan en termes de contenu mais incompatible en termes de processus (p. ex., une interprétation identifiant correctement la peur d'un patient d'agir de façon indépendante peut être exprimée de façon contrôlante), ou vice versa. Une nouvelle analyse de données mises à disposition par Silberschatz et al. nous permet de démontrer que des juges peuvent distinguer de façon fiable le contenu du processus en évaluant la compatibilité avec le plan. Nous démontrons, en outre, qu'une distinction entre compatibilité de contenu et de processus avec le plan peut augmenter la variance expliquée des scores du Vécu du Patient. Une analyse des interprétations dans lesquelles les aspects de compatibilité avec le plan (contenu ou processus) différaient s'avère particulièrement intéressante, mettant en évidence que l'aspect de compatibilité avec le plan le plus hautement corrélé avec le Vécu variait à travers les cas. Ce résultat suggère que de tenir compte des contributions distinctes du contenu et du processus dans la compatibilité avec le plan peut faire avancer notre compréhension de la manière dont un patient spécifique écoute et utilise les interventions du thérapeute.

Resumen

El problema del éxito de las interpretaciones se encuentra entre los más importantes de las investigaciones en psicoterapia. En un estudio señero, Silberschatz, Fretter y Curtis (1986) desarrollaron un método para evaluar la compatibilidad entre una interpretación y la psicología del paciente (el plan inconsciente del paciente) y demostraron que este factor —la compatibilidad del plan—puede contribuir substancialmente al éxito de la interpretación. Sin embargo, estos autores no consideraron la siguiente posibilidad: una interpretación puede ser compatible con el plan en términos de contenido pero incompatible en términos de proceso (v.g.: una interpretación que identifica correctamente un temor del paciente a la independencia puede ser administrada en una forma controladora), o viceversa. En un reanálisis de los datos suministrados por Silberschatz et al., demostramos que, al evaluar la compatibilidad del plan, los evaluadores pueden diferenciar entre contenido y proceso en forma confiable. Demostramos, además, que la distinción entre contenido y proceso en la compatibilidad del plan puede aumentar la variancia de los puntajes de Vivencias del paciente (Patient Experiencing). Es de interés particular el análisis de las interpretaciones en las que difirieron estos aspectos de la compatibilidad del plan (contenido o proceso), porque se vio que el aspecto de la compatibilidad del plan más correlacionado con la Vivencia varió a lo largo de los casos. Este hallazgo sugiere que atender a las diferentes contribuciones de contenido y proceso en la compatibilidad del plan puede ampliar nuestra comprensión de cómo algunos pacientes escuchan y utilizan las intervenciones terapéuticas.

Received August 3, 1996

Revision received March 30, 2000

Accepted April 2, 2000