Patient and Therapist Introject, Interpersonal Process, and Differential Psychotherapy Outcome

William P. Henry
Vanderbilt University

Thomas E. Schacht
James H. Quillen College of Medicine

Hans H. Strupp
Vanderbilt University

The Structural Analysis of Social Behavior (SASB; Benjamin, 1974, 1982, 1984) system was used to study the interpersonal process between patient and therapist in the 3rd session of 14 therapeutic dyads. Dyads were grouped into good and poor outcomes cases (n = 7) on the basis of the amount of change in the patients’ introject as measured by the INTREX Introject Questionnaire (Benjamin, 1983). Strong support was found for the following hypotheses based on interpersonal theory, linking therapists’ introject state, interpersonal process in therapy, and outcome: (a) Poor outcome cases (no introject change) were typified by interpersonal behaviors by the therapist that confirmed a negative patient introject; (b) the number of therapists’ statements that were subtly hostile and controlling was highly correlated with the number of self-blaming statements by the patients; (c) therapists with disaffiliative introjects tended to engage in a much higher level of problematic interpersonal processes that have been associated with poor outcome. Implications for future research and therapist training are discussed.

Stripped of technical jargon, psychotherapy may be seen in one context as simply a structured relationship between two people. Despite the multiplicity of theoretical approaches to the practice of psychotherapy, most major camps recognize the crucial importance of the quality of the therapeutic relationship (Gurman, 1977), and a number of studies have associated an affiliative therapeutic relationship with improved clinical outcomes (Gomes-Schwartz, 1978; O'Malley, Suh, & Strupp, 1983; Sachs, 1983). This consensus, however, belies disputes regarding the exact role of the therapeutic relationship in achieving positive outcomes in therapy.

This issue has received little empirical clarification, in part because of measurement difficulties that arise when one attempts to operationalize the therapeutic relationship. Global scales can rate the overall relationship “climate” with items such as therapist “warmth,” “judgmentalness,” and so forth. Although such global scales have provided useful descriptive accounts that serve in a general sense to distinguish good from poor relationships, they do not explicate exactly what the therapist does to establish the quality of the relationship, and often they are not designed to articulate with specific theories of change. Global scales may thus encourage the view that the relationship is merely a generic or nonspecific background factor forming a context in which specific techniques are applied. The present study was designed to further a more precise understanding of the role of the therapeutic relationship by testing theoretically derived propositions linking therapist interpersonal actions and therapeutic change.

Henry, Schacht, & Strupp (1986) demonstrated a precise method for operationalizing the therapeutic relationship that allowed the therapist's relationship contribution to be seen as a specific technique in itself. Moment-by-moment transactions between therapist and patient were measured with Structural Analysis of Social Behavior (SASB; Benjamin, 1974, 1982, 1984) and compared across good versus poor outcome cases seen by the same therapists. Disaffiliative and complex (mixed-message) interpersonal transactions occurring between patient and therapist were found to be strongly related to subsequent clinical outcome. In terms of the underlying dimensions of SASB, therapists in poor outcome cases were exercising more hostile control and were less prone to grant friendly autonomy, while their patients were responding with a reduced level of affiliative autonomy-taking and greater hostile separation.

This study (Henry, Schacht, & Strupp, 1986) demonstrated the potential importance of dyadic interpersonal process to therapeutic change, but because of limitations imposed by preexisting data, Henry et al. could not evaluate specific hypotheses about change mechanisms. To advance this line of research it is important to refine hypotheses about how interpersonal transactions may produce outcome changes. One important hypothesis derives from the theory of interpersonal introjection.1 The most basic principle of introjection theory is that people learn to treat themselves as they have been treated by others (Sullivan, 1953). The dynamic interaction of these early inter-

1 The term introjection has multiple definitions related to various psychodynamic and developmental theories (see Greenberg & Mitchell, 1983). Our article relies most directly on the theories of Sullivan (1953) and Benjamin (1984).
nalized interpersonal influences forms a hypothesized personality structure termed the "introject," which comprises a relatively stable conscious and unconscious repertoire of ways of treating the self. This repertoire, which internally mirrors treatment received at the hands of early significant others, includes self-appraisals, verbal and motor behaviors directed at the self, and cultivation of various images of the self. Although early experience is thought to be critical, the process of introject formation and the content of the introject are presumably subject to development and change across the lifespan. Individuals who are psychologically healthy, by definition, evidence introjects that are relatively friendly (self-accepting, self-nurturing, self-helping) most of the time. Emotionally disturbed persons, in contrast, tend to have hostile introjects that are recurrently self-critical, self-destructive, or self-neglectful.

Although the descriptive language may vary, a range of theorists from psychodynamic to cognitive-behavioral hypothesize that introjective actions or states are major correlates of affective experience and are important in shaping interpersonal behaviors. For example, self-criticism is associated with shame and sadness, and self-neglect may be associated with boredom. The introject is also hypothesized to play a central role in recurrent maladaptive relationship patterns that underlie symptomatic presentations (Schacht, Binder, & Strupp, 1984; Schacht & Henry, in press). Within these recurrent patterns, introject states are hypothesized to become self-perpetuating through the dynamics of interpersonal complementarity (Benjamin, 1982). For example, a patient's inner self-criticism may lead to defensiveness in interpersonal transactions, implicitly casting the therapist in the role of critic (the accustomed action of important others that originally led to the introject formation). To the extent that the therapist defends himself or herself against this role attribution, such as by disagreeing with the patient, the therapist self-defeatingly becomes a critic and "confirms" the patient's negative expectation (perpetuating the introject state).

Because of its theoretically central role in maintaining problematic affective/interpersonal cycles, a measure of the introject becomes a logical element when evaluating patients' initial status and therapeutic change. As described in the example above, introject theory would predict that the therapist's interpersonal behavior would serve to either perpetuate problematic introject states (leading to poor outcome) or ameliorate them through experiential disconfirmation. If outcome is defined in terms of introject change, then outcome should vary as a function of the therapists' interpersonal behavior. Introject theory also predicts that therapists will tend to treat patients in accord with their own introjects. This will occur by the same complementary dynamics that tend to make introjects self-perpetuating for patients. Thus, therapists whose introjects are self-accepting would be expected to engage their patients in accepting and supportively helpful transactions as compared with therapists with hostile introjects, who would be expected to behave in a relatively more critical or neglectful manner with their patients.

The foregoing theoretical model offers conceptual links between (a) a problem definition (negative introject); (b) a treatment process (differential interpersonal process as mediated by the variable of therapist introject); and (c) an outcome measure (introject change). Our study capitalizes on these conceptual links by using a measurement system (SASB) that allows problems, treatments, and outcomes to be represented in a shared metric that articulates with the underlying theory (cf. Principle of Problem-Treatment-Outcome [P-T-O] Congruence; Schacht, Strupp, & Henry, 1988). The aim of the study is to test process-outcome relationships that would be consistent with the theoretically proposed introject mechanism. Two hypotheses were examined:

1. For patients who begin therapy with disaffiliative introjects, higher levels of therapist disaffiliativeness will be associated with lower levels of introject change or deterioration.
2. Therapists with disaffiliative introjects will treat their patients in a relatively more critical and neglectful manner.

Method

Subjects

Fourteen therapeutic dyads were drawn from the first cohort of the Vanderbilt II psychotherapy research project, a 5-year study of Time-Limited Dynamic Psychotherapy (Strupp & Binder, 1984).1 Dyad selection was structured to produce two equal groups (n = 7) of high-change and low-change cases on the basis of changes in the SASB introject. Participating patients were comparable to a general psychiatric outpatient population, as indicated by intake scores on the Global Severity Index of the SCL-90-R (Derogatis, 1977). The GSI group mean was 48.9, and the two outcome groups were statistically equivalent on the GSI measure at intake. All patients were judged to have significant interpersonal difficulties and qualified for a Diagnostic and Statistical Manual of Mental Disorders, 3rd ed. (DSM-III; American Psychiatric Association, 1980) Axis I or Axis II diagnosis, as determined by a clinician using a computerized rating version of the National Institute of Mental Health (NIMH) Diagnostic Interview Schedule. The mean age of patients was 41.04 (range 24-64 years), 77% were female, 79% had completed at least some college work, 69% had received previous therapy, and all patients and therapists were White. Patients were assigned on a random basis to participating therapists, who were self-described psychodynamically oriented psychiatrists and clinical psychologists in private practice with a minimum of 2 years' full-time postresidency or postinternship experience. Therapies consisted of weekly 50-min meetings for a maximum of 25 sessions and were conducted with no intrusions by the research staff, except to collect written measures of process and outcome.

SASB Measures

The SASB is operationalized as the third of the three interrelated circumplex surfaces of the SASB model. Each circumplex surface presents, in a two-dimensional space, 36 interpersonal behaviors that represent unique combinations of the theoretically "primitive" vectors of affiliation-disaffiliation and independence-interdependence. Interpersonally complementary behaviors are represented as homologous points across the surfaces. Each SASB surface defines a particular perspective or focus of interpersonal transactions. Surface 1 involves focus on another person (transitive action). Surface 2 involves focus on the self (intransitive states). Surface 3 represents intrapsychic actions that result when focus on the other (Surface 1) is directed inward on the self.

---

1 Fourteen was the maximum number of dyads with clearcut good and poor outcomes that were available for analysis at the time the study was conducted.
The reader is directed to Benjamin (1974, 1982) for a more complete description of the SASB system. Figure 1 presents an alternate version of the SASB model in which the 36 behaviors for each circumplex are collapsed into eight clusters. Data for the present study are presented in terms of these clusters.

**Initial status and outcome measurement.** The SASB system affords two methods for measuring the introject. The first method involves direct ratings of observed introjective behaviors. For example, a statement such as “I’m an awful person for thinking that” would be coded as a self-critical introject (Surface 3, Cluster 6, or 3-6). Direct codings of transactions were employed in the present study as measures of treatment process. The second measurement method, used in the present study as the measure of initial state and outcome, uses the INTREX Introject Questionnaire (Benjamin, 1983). This well-validated instrument comprises 36 items, each corresponding to one of the 36 introjective behaviors that compose the introject circumplex of the SASB model. Sample items are given in Table 1. Self-ratings are made on a scale of 0-100 according to the degree to which the rater judges an item to be descriptive of his or her behavior. Each item is rated twice, once with an instruction to “describe yourself at your best” and once with an instruction to “describe yourself at your worst.” The result is two measures, “introject at best” and “introject at worst.” Multiple administrations of the INTREX Questionnaire allow changes in the patient’s introject at best and worst to be graphed over time.

INTREX introject ratings are scored by a computer program (Benjamin, 1983) that weights each rating in relation to its position on the basic circumplex dimensions of control versus freedom and attack versus affiliation. The result is a measure of central tendency that summarizes the overall introjective stance of the rater. This measure of central tendency is expressed in terms of two coefficients, corresponding to the circumplex axes, that define a point on the circumplex surface. The first coefficient, termed the attack coefficient, expresses the central...
tendency of ratings on the basic dimension of self-affiliation versus self-attack (self-love vs. self-hate). The second coefficient, termed the control coefficient, expresses the central tendency of the ratings on the basic dimension of self-freeing versus self-control. Values for these coefficients are expressed in a range from $-1,000$ to $+1,000$.

Patients were assigned to high-change (good) and low-change (poor) outcome groups according to the amount and direction of change in their self-reported introject at worst as measured by the INTREX Introject Questionnaire administered at intake; Sessions 3, 8, 16; and termination. All patients in the current study began therapy with their introject at worst in either Quadrant 2 or 3 of the introject circumplex (Figure 1, Surface 1). This reflects a hostile and critical or neglectful stance toward the self. There was no statistically significant difference between the outcome groups at intake in the amount of self-hostility (as represented by the attack pattern coefficient) at best or worst. Therapists also completed INTREX Introject Questionnaires rated at best and worst on themselves prior to meeting their patients.

Indications of positive change were defined as (a) clinically significant movement from left to right on the circumplex, indicating greater self-affiliation (approximately 500 points on the $-1,000$ to $+1,000$ axes); (b) movement from Quadrant 2 to 3. (Quadrant 2 is considered the most pathological state, as it represents both self-hatred and self-freeing [neglect], which is hypothesized to be a potentially suicidal introject profile.) Thus movement to Quadrant 3, reflecting greater self-control and self-management, is deemed positive therapeutic change.

Indications of low change were defined as (a) movement from Quadrant 3 to 2 (away from self-control), which represents a change toward greater self-neglect; (b) significant movement from right to left indicating increased self-hostility; (c) no change, including no improvement in introject at best or worst.

To explore the validity of this procedure, the two groups formed on the basis of introject change were compared on more traditional outcome measures. Pre-post residual change scores (partialing out variance in change scores due to initial status) were calculated on the Global Assessment Scale (GAS) as rated by an independent clinician and the SCL-90-R Global Severity Index. In addition, the Global Outcome Ratings made separately by the patient and therapist after therapy ($-5$ to $+5$ direct change rating) were examined. Independent groups $t$ tests were all significant ($p < .05$) with differences in the expected direction. Thus measures of level of functioning, symptom status, and perceived change from three sources were consistent in supporting the division of the patient sample into positive and negative outcome on the basis of introject change.

## Results

### Interpersonal Process and Differential Outcome

Before statistical analysis, the observed frequencies in each of the 16 possible SASB focus/cluster combinations were tallied for both patient and therapist in each case. These frequencies were then weighted to the respective patient or therapist grand mean for number of codings to remove the artifact of differing proportions of turns at talk between the participants in different dyads. Such weighted frequencies, the basic data studied, allow for meaningful cross-case comparison.

Initially, two $2 \times 2 \times 2$ (Outcome $\times$ Focus $\times$ Cluster) ANOVAs with one between and two within factors were performed (one analysis for patient communications and one for therapist communications), using weighted cluster frequencies as the dependent variable. Simple effects tests for differences in cluster means between the two outcome groups were accomplished using between-groups $t$ tests. In several instances the Mann-Whitney $U$ statistic was substituted as a result of significant differences in within-group variance.

As expected, therapists spent more time focusing on other ($F = 354; p < .001$), whereas patients spent more time focusing on self ($F = 2642; p < .001$), and the eight SASB clusters were utilized differently depending on outcome group ($F = 4.8; p < .001$). Table 2 contains the results of the cluster comparison tests. Patients in the poor outcome group were significantly more watching and managing toward the therapists, more as-

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Without concern I just let myself be free to turn into whatever I will.</td>
</tr>
<tr>
<td>2</td>
<td>Knowing both my faults and strong points I comfortably let myself be &quot;as is.&quot;</td>
</tr>
<tr>
<td>3</td>
<td>I like myself very much and feel very good when I have a chance to be with myself.</td>
</tr>
<tr>
<td>4</td>
<td>I practice and work on developing worthwhile skills, ways of being.</td>
</tr>
<tr>
<td>5</td>
<td>I have a habit of keeping very tight control over myself.</td>
</tr>
<tr>
<td>6</td>
<td>I accuse and blame myself until I feel guilty, bad, and ashamed.</td>
</tr>
<tr>
<td>7</td>
<td>I harshly punish myself, take it out on myself.</td>
</tr>
</tbody>
</table>
| 8       | Instead of getting around to doing what I really need to do for myself, I let myself go and just dream day.

---

1. Introject-at-best ratings were not the variable of primary interest in our analysis because they tended to be idealized, resulting in a restricted range of ratings.

2. This figure is considered conservative and was arrived at in consultation with L. S. Benjamin, who has extensive clinical experience with the INTREX Questionnaires.

3. Each part of a complex communication was counted as one occurrence of the appropriate cluster. For instance, a 2-2/2-6 was counted as one 2-2 and one 2-6.

4. After we consulted several statisticians, adjustment for familywise error rate was deemed unnecessary in this case because of the relatively small number of simple effects tests performed. At the .05 level of significance, no results (less than one) would be expected by chance alone for each ANOVA.
serving and separating, sulking and appeasing, and walling off and avoiding, while being less disclosing and expressing. Therapists in the poor outcome group were significantly more belittling and blaming and ignoring and neglecting. Thus, replicating our previous study, patients and therapists in the low-change group demonstrated significantly higher levels of disaffiliative behavior. Additional analyses were performed yielding the following results:

1. A strong correlation was observed between the number of therapist statements that were hostile and controlling (Focus-Other, Quadrant 3) and the number of their patient’s statements that were self-blaming and critical (Introject Quadrant 3), $r = 0.53$ ($p < .05$).

2. Both patients and therapists in the low introject change group were found to exhibit significantly higher levels of complex communications (see Table 2). These included primarily mixed affiliation statements (simultaneously affiliative and hostile) and mixed focus statements (focus on self and other simultaneously). The relative percentage of mixed message thought units for good versus poor outcome cases was 3% versus 13% for the patients ($p < .04$) and 6% versus 26% for the therapists ($p < .01$).

3. The percentage of thought units receiving at least one disaffiliative code was calculated for the patient and the therapist, yielding an index of hostile communication. Both patients (16% vs. 2%) and therapists (21% vs. 3%) in the poor outcome group acted with significantly greater hostility ($p < .04$).

4. Relative to their therapists, patients in the poor outcome group spoke less than patients in the good outcome group (69% vs. 79%, $p < .05$, two-tailed).

**Therapist Introject and Differential Interpersonal Process**

As predicted, therapists whose pretherapy self-ratings indicated relatively greater hostility toward themselves were also more likely to be coded as treating their patients in a disaffiliative manner. Therapists whose introjects rated at worst were friendly (Quadrants 1 and 4) were given hostile codes in 5.6% of the coded transcript thought units. In contrast, therapists who reported hostile introjects at worst (Quadrants 2 and 3) were given hostile codes 17.7% of the time (Mann-Whitney $U$, $p < .05$).

**Discussion**

Coupled with earlier findings (Henry et al., 1986), this study provides strong evidence for the importance of interpersonal transactions early in therapy. It should be clearly noted that the present design does not offer proof of the hypothesized underlying mechanisms, but it does offer a range of data consistent with predictions based on interpersonal theory. There was a significant relationship between therapists’ interpersonal behaviors and the ways in which the patients acted toward themselves. Patients whose introjects showed marked improvements
engaged in therapeutic interactions that were almost completely devoid of disaffiliative therapist behaviors, whereas therapists of those patients showing no introject change (or deterioration) engaged in hostile blaming, ignoring, and separating sequences.7 This relationship between interpersonal process and differential intrapsychic changes suggests one important specific role of the therapeutic relationship in promoting positive outcome.

Related predictions were also supported. Carson (1969), integrating the theories of Sullivan and others (Thibaut & Kelly, 1959; Secord & Backman, 1961), stated that one's interpersonal behavior is unconsciously designed to evoke responses from others that validate a stable self-image (even a negative self-concept, see Anchin, 1982). The results indicate that there was a high degree of correspondence between patient self-blaming statements and therapist statements subtly blaming the patient \( r = .53 \). Finally, in what is considered a most significant finding, the therapists' own introjects were found to be systematically related to their in-session process with patients.8 These results may have important theoretical, methodological, and training implications.

From a methodological standpoint, the study demonstrates the utility of defining problems, therapeutic process, and outcome in similar theoretically linked metrics. In a broader sense, it demonstrates ties between specific therapist behaviors and clinical change, a line of research with an unfortunate history of failure (Gurman, 1977; Orlinsky & Howard, 1986). Although frequency counts were used as the unit of analysis, these units represented momentary dyadic states, not aggregation of unilateral acts. This is consistent with an increasing call for context-sensitive approaches to process research (Greenberg, 1986). More generally, it appears that interpersonal process analysis with SASB may provide an important contextual backdrop when combined with other descriptive process measures. For instance, it would be valuable to study the effect on interpersonal process of various types of "technical" interventions, such as transference interpretations, with different types of patients during different phases of therapy. There has been recent suggestive evidence that such interpretations often foster resistance and poor process (Foote, McCullough, Winston, Pollack, & Laikin, 1988; Marmar, Weiss, & Gaston, 1988). Perhaps it is not the nature of the interventions that is responsible for these findings, but the interpersonal context in which they are often carried out. Wile (1984), writing on the accusatory nature of many common interpretations, states:

Certain interpretations commonly made in psychoanalysis and psychodynamic therapy are accusatory. Therapists appear to make them not because they are hostile or insensitive, but because of the dictates of their theory. Clients are seen as gratifying infantile impulses, being defensive, having developmental defects, and resisting therapy. . . . Even therapists who reject the psychoanalytic model may nevertheless view clients as dependent, narcissistic, manipulative, and resistant, and as refusing to grow up and face the responsibilities of adulthood. Therapists who conceptualize people in these ways may have a hard time making interpretations that do not communicate at least some element of this pejorative view (p. 353). . . . The major difficulty with accusatory interpretations, in conclusion, is that clients are already accusing themselves, often for the very thing for which the traditional therapist is accusing them, and that this—self-criticism—is the problem. (p. 363)

The high correlation observed in the present study between therapist-blaming and patient self-blaming would seem to bear out Wile's point.

Our data suggest an important overriding principle: Whereas the absence of a negative interpersonal process may not be sufficient for therapeutic change, the presence of even relatively low levels of negative therapist behavior may be sufficient to prevent change. Why should this be so? One explanation rests on a cognitive understanding of the introject. Introjects may be construed as overlearned beliefs or stereotypes about the self. By definition, a stereotype is a belief that is held rigidly and is not easily opened to question or disconfirmation. Very little apparently affirmative "interpersonal evidence" may be required to prove a stereotype in the mind of the believer. To promote change, the theory predicts that therapists' behavior should rather consistently disconfirm patients' negative self-expectations.10 Similar therapeutic prescriptions appear in Beier and Young's (1984) concept of "asociative response" and in Weiss and Sampson's (1986) assertion that therapists must repeatedly pass patients' interpersonal "tests."

Conclusion

Apparently, even well-trained professional therapists are surprisingly vulnerable to engaging in potentially destructive interpersonal process (in the most extreme case observed, the therapists' communications were judged to be disaffiliative 56% of the time, with 65% of the thought units found to contain complex messages). Perhaps this vulnerability should not be so surprising when one considers the strong pull of multiple potential determinants of such process. Patients' preexisting maladaptive "evoking styles" (Kiesler, 1982) often tend to elicit poor process according to the rules of interpersonal complementarity. The therapist's own introjective state may also contribute to problematic process. In addition, as Wile noted, theoretical heuristics may themselves inadvertently encourage disaffiliative interpersonal transactions. It seems clear that

---

1. It is interesting to note that when our two studies were combined, in 11 good outcome cases, not a single instance of therapist blaming and belittling (Cluster 6) was observed.

2. It should be noted that this relationship is not a given, as patients in the good outcome group also made self-blaming statements (although at a lower frequency), but there were no instances in which these statements were complemented by therapist blaming behaviors.

3. This finding cannot be taken to mean that all therapists with "disaffiliative introjects" will display hostility toward all patients. It does suggest that some therapists, as a function of their own interpersonal history, may be more vulnerable to engaging in countertherapeutic interpersonal processes with patients who evoke such negative tendencies.

4. It should be noted that to perfectly disconfirm patients' expectations is quite likely an impossible task. Additional research should address the question of how therapists who find themselves in a disaffiliative transaction can best repair the situation. We hypothesize that therapeutic metacommunication plays a crucial role in this task.
traditional training methods have not adequately prepared many therapists to expertly perceive and respond to the potential interpersonal process meanings and effects of their interventions. Future training programs, regardless of theoretical orientation, may be designed to directly address these fundamental interpersonal skills.

Finally, our results potentially point to an underlying change process operative across therapies with different surface structures. In speaking of the fundamentally human and interpersonal nature of psychotherapy, Safran and Segal (1990) caution, "relevant theory must clarify the process through which this human encounter brings about change." We suggest that the role of the therapeutic relationship in fostering therapeutic change through its effect on the patient's introject structure constitutes a common change mechanism of central importance. When the relationship is seen in this light, the distinction between so-called nonspecific and specific change factors disappears.

References


Received August 17, 1989
Revision received April 2, 1990
Accepted April 23, 1990