

This article was downloaded by: [Grande, Tilman]

On: 6 July 2009

Access details: Access Details: [subscription number 912941492]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Psychotherapy Research

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713663589>

### Structural change as a predictor of long-term follow-up outcome

Tilman Grande <sup>a</sup>; Reiner Dilg <sup>b</sup>; Thorsten Jakobsen <sup>a</sup>; Wolfram Keller <sup>b</sup>; Bärbel Krawietz <sup>a</sup>; Monika Langer <sup>a</sup>; Claudia Oberbracht <sup>a</sup>; Sabine Stehle <sup>b</sup>; Margret Stennes <sup>b</sup>; Gerd Rudolf <sup>b</sup>

<sup>a</sup> Department of Psychosomatic and General Clinical Medicine, University of Heidelberg, Germany <sup>b</sup>

Department of Internal and Psychosomatic Medicine, Theodor Wenzel Werk Berlin, Germany

Online Publication Date: 01 May 2009

**To cite this Article** Grande, Tilman, Dilg, Reiner, Jakobsen, Thorsten, Keller, Wolfram, Krawietz, Bärbel, Langer, Monika, Oberbracht, Claudia, Stehle, Sabine, Stennes, Margret and Rudolf, Gerd(2009)'Structural change as a predictor of long-term follow-up outcome',*Psychotherapy Research*,19:3,344 — 357

**To link to this Article:** DOI: 10.1080/10503300902914147

**URL:** <http://dx.doi.org/10.1080/10503300902914147>

## PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## Structural change as a predictor of long-term follow-up outcome

TILMAN GRANDE<sup>1</sup>, REINER DILG<sup>2</sup>, THORSTEN JAKOBSEN<sup>1</sup>, WOLFRAM KELLER<sup>2</sup>,  
BÄRBEL KRAWIETZ<sup>1</sup>, MONIKA LANGER<sup>1</sup>, CLAUDIA OBERBRACHT<sup>1</sup>,  
SABINE STEHLE<sup>2</sup>, MARGRET STENNES<sup>2</sup>, & GERD RUDOLF<sup>2</sup>

<sup>1</sup>Department of Psychosomatic and General Clinical Medicine, University of Heidelberg, Germany & <sup>2</sup>Department of Internal and Psychosomatic Medicine, Theodor Wenzel Werk Berlin, Germany

(Received 22 April 2008; revised 11 March 2009; accepted 16 March 2009)

### Abstract

Based on data from psychoanalytic long-term psychotherapies, the predictive value of three measures of pre–post change for retrospective patient assessments of outcome at 1-year and 3-year follow-up was investigated. Pre–post changes were measured using the Global Severity Index (GSI), the Inventory of Interpersonal Problems (IIP) total score, and the Heidelberg Structural Change Scale (HSCS). In line with psychoanalytic theory, it was assumed that structural changes cause especially persistent changes and would, therefore, be most suitable to predict the follow-up criterion. This expectation was confirmed: Pre–post changes in GSI and IIP were only weakly associated with assessments at 1-year follow-up and not at all with assessments at 3-year follow-up. In contrast, correlations between changes in HSCS and outcome assessments were highly significant at both occasions.

**Keywords:** long-term psychotherapy; outcome research; psychoanalytic/psychodynamic therapy

Follow-up studies are generally conducted to establish whether the effects of psychotherapy remain stable following the completion of therapy. Of primary interest in such investigations is the effectiveness or the persistence of effects with respect to a specific form of treatment compared with another form or a control condition. These studies provide evidence concerning whether the therapeutic effect of a treatment (e.g., symptom reduction) remains constant or is even enhanced up to a specified follow-up measurement occasion (Kendall, Holmbeck, & Verduin, 2004; Lambert & Ogles, 2004).

In the present study, we adopt a different approach in investigating the long-term effects of therapy outcomes. We address the following question: Which specific type of change should occur by the end of therapy in order to ensure that broad positive effects, which influence many central life domains, are later (at a follow-up) found? Might various forms of pre–post change be associated with such effects, according to whether these are measured at the end of therapy or after a longer interval of time has elapsed following therapy completion? The following example illustrates this point: Let us assume that an investigation measures changes in

symptom distress and interpersonal problems from the onset to the end of therapy. An outcome criterion that reflects treatment effects in various life domains is identically applied at the end of therapy and at a follow-up. It is found that, although change in symptom distress and interpersonal problems both highly correlate with the criterion at the end of therapy, only one of the two measures (e.g., changes in interpersonal problems) continues to significantly correlate with assessments at 3-year follow-up. In comparison with changes in symptom distress, this result may suggest that pre–post changes in interpersonal problems represent a more important outcome in view of its association with long-term, extensive positive changes in the life of the patient. In a further step, this finding might be used to infer the type of change on which therapeutic work should focus.

Psychoanalysis postulates a special form of change that is assumed to account for particularly persistent effects: “structural change.” Structural changes are differentiated from more superficial changes, under which, for example, symptom reduction is subsumed. The term “structure” refers to the temporally stable organisation of the personality and the

---

Correspondence concerning this article should be addressed to Tilman Grande, Psychosomatische Klinik, University of Heidelberg, Thibautstrasse 2, Heidelberg D-69115, Germany. E-mail: tilman\_grande@med.uni-heidelberg.de

habitual patterns that individuals adopt in an attempt to resolve their unconscious conflicts (Rapaport, 1960). It is assumed that changes at this deeper level of the personality are essential in attaining persistent therapeutic effects at all levels (e.g., also at the symptomatic level). Psychoanalytic treatment and the techniques used within psychoanalysis (e.g., encouragement of regression, interpretation) aim to promote such structural changes (Kernberg, 1991, 1999; Moore & Fine, 1990). Investigations suggest that changes of this type do not represent an exclusive characteristic of psychoanalysis but can rather also be achieved within other forms of psychotherapy, such as expressive or supportive approaches (Wallerstein, 1988). When considering the pivotal importance of structural change for the self-image of psychoanalytic therapy, it is astounding that, for a considerable length of time, this concept has remained largely undefined, has scarcely been empirically investigated, and has only recently been subject to operationalisation attempts. These attempts include the Karolinska Psychodynamic Profile developed by Weinryb and Rössel (1991); Scales of Psychological Capacities by Wallerstein and coworkers (DeWitt, Hartley, Rosenberg, Zilberg, & Wallerstein, 1991; Wallerstein, 1996); Reflective Functioning Scale designed by Fonagy, Target, Steele, and Steele (1998); and Heidelberg Structural Change Scale (HSCS; Rudolf, Grande, & Oberbracht, 2000), which is based on the Operationalized Psychodynamic Diagnostic system (OPD Task Force, 2001, 2008). Despite substantial differences between these instruments, a common feature is their reliance on complex clinical ratings on patients and their changes, which can only be conducted by well-trained raters; patients are not considered able to perform self-ratings of structural changes.

With the exception of the HSCS, the just-listed instruments are conceptually designed to measure change as a reduction in unfavourable personality styles or as a dissolution of inner inhibitions or other constraints. Change is thus implicitly defined as an abatement or eradication of pathology. This definition is, however, conceptually problematic when it comes to measuring change in the context of psychoanalytic therapy. Follow-up studies (Leuzinger-Bohleber, 2002; Pfeffer, 1959; Schlesinger & Robbins, 1975) have repeatedly shown that patients remain susceptible with respect to their central conflicts for a long time after successful courses of therapy and that they transiently react in a neurotic manner when conflict-laden topics are touched upon. What is fundamentally changed is rather the patients' ability to deal with such situations in a regulatory fashion: Following therapy, patients are

able to find healthy solutions instead of the neurotic solutions that were previously selected (Nedelmann, 1981; Pfeffer, 1959; Schlesinger & Robbins, 1975). These observations suggest that changes within psychoanalytic treatment should be conceptualized as changes in dealing with conflictual tendencies and vulnerabilities rather than their elimination. Adaptive handling of critical situations becomes possible when patients gain cognitive as well as emotion-based insight into their own complex of problems (Thomä & Kächele, 1987). As discussed in detail later, this form of change is captured by the HSCS, which measures therapy-based changes in dealing with individually defined problem areas (Grande, Rudolf, Oberbracht, & Pauli-Magnus, 2003; Rudolf et al., 2000).

In psychoanalytic literature, it is argued that, in contrast to purely symptomatic changes or changes in overt behavioural patterns, structural changes broadly impact many life domains and are associated with a change of self or the experience of self (Wallerstein, 1965). Changes of this quality are not easily captured using, for example, pre-post comparisons, because it is not only certain aspects of the patient (e.g., symptoms, behaviour patterns, relationship problems) that are subject to change but the entire reference system within which the patient evaluates his or her experience and behaviour. For instance, according to Sandler and Sandler (1983), this occurs when the patient attains reconciliation with previously unaccepted parts of the self and in doing so is able to temper judgments on the self and others. In some circumstances, it may be possible that a problem (e.g., in the interpersonal domain) can, therefore, be solved without the overt changes that the patient desired at the onset of therapy given changes in his or her internal judgments. In this case, the patient would nonetheless retrospectively state that his or her interpersonal difficulties have improved as a result of therapy. Similarly, a patient may also retrospectively note favourable changes in life domains in which no problems had been perceived prior to therapy onset (which are, however, critically identified with hindsight). Retrospective evaluations of change thus reflect shifts within a patient's internal reference system in addition to manifest changes in symptoms and behaviour. In our view, it is thus worth considering whether retrospective evaluations might for that very reason represent an option in assessing more fundamental therapeutic changes from the perspective of the patient.

Retrospective evaluations are viewed critically in the research literature. Pre-post measures or rather pre-post comparisons (Hill & Lambert, 2004; Lambert & Hill, 1994) are considered to be a standard for measurement of success. Critics

importantly point out that numerous investigations have repeatedly demonstrated only a weak correlation between retrospective evaluations and prospectively conducted pre-post measurements. As indicated by Michalak, Kosfelder, Meyer, and Schulte (2003), an implicit assumption of these assessments is that patients retrospectively calculate the difference between their conditions pre- and posttherapy. Their assessments are thus to be seen as “perceived change” (Lunnen & Ogles, 1998; Lunnen, Ogles, & Pappas, 2008; Pekarik & Wolff, 1996). In this case, the lack of correlation with objectively measured pre-post changes shows that patients’ retrospective perceptions are invalid. However, in contrast to this line of reasoning, Michalak et al. and Schulte (2008) have pointed out that patients do not necessarily expect a certain level of symptom change as a result of their therapy but rather the attainment of personal goals, which can vary considerably according to the individual. In retrospective outcome evaluations, symptom changes are, therefore, not objectively (i.e., as a measure of pre-post differences) evaluated but rather are assessed in terms of the attainment of a personal goal. In line with this argumentation and with reference to the considerations discussed previously, we add that fundamental therapeutic changes can further result in a modification of these individual goals and, in turn, the patient’s evaluation criteria. We, therefore, join Michalak et al. and Schulte in advocating that retrospective evaluations of change be entitled for use as an instrument to assess outcome alongside pre-post comparisons. Such an instrument was applied in the current study to assess the short- and long-term effects of treatment in various life domains.

According to psychoanalytic principles, the process of change does not conclude with the completion of therapy. In the “postanalytic phase” (Thomä & Kächele, 1987), further important development steps occur, for example, the final resolution of the transference relationship and identification with the function of the analyst, which is associated with the forming of self-analytic skills (Ticho, 1967). Although the end of therapy to a certain extent marks a completion of the work of the analyst and in favourable cases the achievement of structural changes, many effects do not become apparent until later, when the patient has attained a higher level of autonomy and has acquired methods of independent self-regulation on the basis of his or her newly gained insights. In light of this delayed manifestation of therapeutic effects, psychoanalytic researchers call for a distinction between treatment outcomes measured at the end of therapy and those measured later (Leuzinger-Bohleber, Stuhr, Rüger, & Beutel, 2001;

Thomä & Kächele, 1987; Wallerstein, 2001). An interval of between 2 and 5 years is recommended as an adequate time frame for a follow-up at which persistent therapeutic effects in the life of the patient can be detected (Wallerstein, 2001). In line with this view, we assume in the present study that structural changes (pre-post) in particular predict therapeutic effects, which are measured after a longer posttherapy interval. This assumption finds a certain amount of empirical support in studies that have demonstrated a particularly long-term effect of “therapeutic insight” (cf. Connolly Gibbons, Crits-Christoph, Barber, & Schamberger, 2007).

The present investigation is based on data collected in the Heidelberg-Berlin Study (Grande et al., 2006; Grande, Rudolf, Oberbracht, Jakobsen, & Keller, 2004; Rudolf et al., 2002), which investigated the differential effects of two forms of psychodynamic therapy. Rather than focusing on these differential treatment effects, however, the current investigation adopted the research approach outlined at the outset of the introduction to determine which pre-post changes (i.e., changes from the onset to the end of therapy) in the total patient group best predicted retrospective patient assessments at the end of therapy and at 1-year and 3-year follow-up assessments. It was assumed that structural change would be a better predictor of the criterion compared with changes in symptoms or interpersonal problems and that this would apply regardless of treatment group. However, in order to examine a potential influence of group on the association between predictors and criterion, interactions between pre-post measures and treatment form were nonetheless subject to statistical testing.

## Method

### Participants

The original study was conducted as a conjoint investigation in Heidelberg and Berlin. Design and results are described in detail in Grande et al. (2006). Psychoanalysts in private practice were requested to include consecutive patients seeking therapeutic aid. Patients with psychotic disorders and those younger than 18 years were excluded. We initially included 62 cases in which patients completed their therapy as planned and also participated in the study to the very end. There were nine dropouts. Five other patients terminated their study participation during the course of therapy while continuing with their treatment.

The study was originally designed to compare two therapeutic approaches, namely psychoanalytic (PA) and psychodynamic (PD) treatment (cf. Grande

et al., 2006). Because of the nature of the original study design and the fact that the initially planned therapy method (PA or PD) was so extensively modified by the respective therapist in the course of treatment that it was no longer possible to allocate them to their original group, three of the 62 cases were excluded from analysis, leaving a sample of 59: 32 in the PA group and 27 in the PD group. These patients formed the basis of our analyses. Of these, 55 patients (93.2%; 29 PA and 26 PD) participated in the 1-year follow-up and 53 (89.8%; 29 PA and 24 PD) in the 3-year follow-up.

The average age of the 59 patients was 37.3 years ( $SD=9.4$ ); 39 (66.1%) were women and 20 (33.9%) were men. Thirty-two (54.2%) patients were high school graduates; the remaining 27 (45.8%) left school at an earlier stage. Syndrome diagnosis was performed by therapists using the International Classification of Diseases (ICD-10; Dilling, Mombour, Schmidt, & Schulte-Markwort, 1994). Despite potential doubts concerning the validity of these diagnoses, this approach was selected in order to forgo a detailed exploration of symptomatology by external raters. During the planning of the study together with the psychoanalysts, this procedure was instrumental in substantially enhancing study acceptance. All analysts were provided with ICD-10 research criteria (Dilling et al., 1994) and were also given the opportunity to discuss and clarify respective cases with project representatives in instances of diagnostic uncertainty. Analysts took frequent advantage of this option. Personality disorders were diagnosed in accordance with ICD-10 (F60, F61) by independent raters on the basis of interviews (carried out in line with OPD guidelines; see later). These raters were highly experienced in the use of the ICD-10. Because the narcissistic personality disorder is not included in the ICD-10, this diagnosis was undertaken in accordance with *Diagnostic and Statistical Manual of Mental Disorders* (fourth edition [DSM-IV]; American Psychiatric Association, 1994) criteria (and encoded as F60.81). The most frequent ICD-10 diagnoses were depressive disorders (F32–34: 66.1%), anxiety disorders (F41: 40.7%), and somatoform disorders (F45: 37.3%), followed by compulsive disorders (F42: 22.0%), sexual dysfunctions (F52: 22.0%), adjustment disorders (F43: 18.6%), and eating disorders (F50: 18.6%). Multiple diagnoses were allowed. Of the total, 31 (52.5%) patients were diagnosed with a personality disorder, most frequently narcissistic ( $n=11$ , F60.81: 18.6%) or borderline ( $n=6$ , F60.31: 10.2%). (Further personality disorders included emotionally unstable personality disorder, impulsive type, F60.30, dependent personality disorder, F60.7, and histrionic

personality disorder, F60.4:  $n=2$ ; paranoid personality disorder, F60.0, anankastic personality disorder, F60.5, and anxious avoidant personality disorder, F60.6:  $n=1$ ; dual diagnosis F60.1 and F60.81,  $n=1$ ; unspecified personality disorders,  $n=4$ ). An average of 2.5 clinical diagnoses (comparable with Axis I in DSM) and three diagnoses including personality disorders were made per patient.

The Global Severity Index (GSI) of the Symptom Checklist-90-Revised (SCL-90-R; Derogatis, 1994; German version: Franke, 2002) was used to measure current overall distress. The mean value at the onset of therapy was 1.06 ( $SD=0.59$ ). In comparison, Brockmann, Schlüter, Brodbeck, and Eckert (2002) reported a mean GSI value of 0.92 ( $SD=0.54$ ) for an unselected sample of 31 patients at the onset of psychoanalytically oriented outpatient psychotherapy, and Schauenburg and Strack (1999) found a mean value of 1.22 ( $SD=0.65$ ) in a mixed group of 410 patients from psychoanalytic practices.

For the total score of the Inventory of Interpersonal Problems (IIP; Horowitz, Strauss, & Kordy, 2000), we found a mean value of 1.62 ( $SD=0.49$ ) at therapy onset. In comparison, Brockmann et al. (2002) reported a mean value of 1.69 ( $SD=0.43$ ) at the start of therapy for their unselected patient sample. In a controlled study with a sample of 63 patients receiving outpatient psychoanalysis for a depressive disorder, Huber, Henrich, and Klug (2007) found a mean IIP total value of 1.81 ( $SD=0.38$ ). In sum, the degree of impairment shown by our patients was thus representative of that generally found under naturalistic conditions in psychoanalytically oriented outpatient therapies in Germany.

### Therapists

To qualify for participation in the study, therapists were required to have completed psychoanalytic training at an institute recognized by the German Association for Psychoanalysis, Psychotherapy, Psychosomatics, and Depth Psychology (DGPT), the umbrella organisation for psychoanalytic therapy schools in Germany. A course of training as stipulated by the statutes of the DGPT qualifies and entitles therapists to recover the costs of PA and PD therapies from statutory German health insurance schemes. The 59 courses of therapy were conducted by 48 analysts from Heidelberg and Berlin. Some therapists ( $n=14$ ) contributed one PA and one PD case to the study and others either one PA or one PD case. The mean age of therapists was 51.8 years ( $SD=6.7$ ); 29 (60.4%) were women and 19 (39.6%) men. Twenty-six (54.2%) were psychologists and 22 (45.8%) were physicians (psychiatrists). Average professional experience as psychotherapist

was 20 years ( $SD = 7.59$ ; range = 5–40 years), and approximately half had completed their psychoanalytic training at least 7 years ( $SD = 7.35$ ; range = 2–31 years) before the start of the project. Accordingly, all analysts involved in the study were well-trained and experienced psychotherapists.

## Treatment

Differences between the two psychodynamic therapeutic approaches PA and PD did not form the focal point of this investigation. Therapeutic form was, however, considered as a potential moderator. PA was primarily (for more than half of therapy duration) conducted with the patient in a lying position, with a frequency of at least three sessions per week and a total of at least 150 sessions. PD was for the most part conducted with the patient in a sitting position, with a frequency of one session per week and a total of between 25 and 100 sessions. Given its connection with the therapeutic techniques used and the objectives associated with the respective treatment approach, session frequency represented a crucial defining factor. With an average duration of 44.2 months ( $SD = 14.3$ ,  $Mdn = 43.8$ ), PAs lasted almost twice as long as PDs, which had an average duration of 24.2 months ( $SD = 8.5$ ,  $Mdn = 23.1$ ). In the PA group, the number of sessions ( $M = 310$ ,  $SD = 102.9$ ,  $Mdn = 300$ ) was more than four times higher than in the PD group ( $M = 71.1$ ,  $SD = 25.5$ ,  $Mdn = 75$ ). Variances in both groups were, however, substantial, reflecting the wide range of session numbers.

As mentioned, PA and PD are associated with different therapeutic techniques and objectives. Therapists used an item checklist every 3 months to indicate their present therapeutic aims and the treatment techniques used. In line with the standard commentary on the German guidelines for psychotherapy (Rüger, Dahm, & Kallinke, 2003), supportive and focal interventions and objectives were evaluated as being characteristic of PD (items: clarify, advise, relieve stress, encourage, structure, focus-oriented work on personality problems), whereas regression and transference-oriented approaches were classified as being characteristic of PA (items: encouragement of/work on transference, admission/encouragement of regressive processes, work on dreams, unrestricted and extensive work on personality problems). In each case, these items were summated to form a PD or a PA score, which was then divided by the number of items (Grande et al., 2006). Scores thus ranged from 0 to 1, with a maximum score of 1 indicating that all items typical of one of the treatment forms had been checked. Scores were averaged across all assessment occa-

sions. In this way, differences between the PD and PA groups were established. The average PA score was .50 ( $SD = .19$ ) in the PA group and .23 ( $SD = .17$ ) in the PD group. This difference was significant,  $t(56) = 5.747$ ,  $p < .001$ , one-tailed. The average PD score was .22 ( $SD = .15$ ) in the PA group and .36 ( $SD = .16$ ) in the PD group. This was once again significant,  $t(56) = 3.64$ ,  $p < .01$ , one-tailed. Effect sizes (Cohen, 1977) were 1.53 for the PA score and 0.96 for the PD score.

## Measures

Structural changes were measured using a two-step approach. In the first step, core problem areas were defined for each patient based on the Operationalized Psychodynamic Diagnostics (OPD). The second step involved an assessment of change for each patient over the course of therapy with respect to the identified problem areas. Outcome measures were thus individualized. The OPD (OPD Task Force, 2001) comprises three psychodynamic axes that are relevant in this context (cf. Table I): The relationship axis captures the dominant dysfunctional relationship patterns displayed by a patient. The conflict axis assesses the presence and degree of intensity of eight types of conflict. The structure axis defines patients' functional levels with regard to six structural capacities or vulnerabilities. (Note: Given the similar-sounding terms "structure" and "structural change," it is necessary to clearly distinguish between the two: Whereas "structure" in the OPD refers to psychological capacities or deficits, "structural change" in the context of the psychoanalytic discussion denotes a basic form of personality modification with respect to relationship patterns, unconscious conflicts, and patients' structural features in the sense of the OPD.)

The validity of the OPD has been investigated in a large number of studies, the results of which have been summarized and discussed by Cierpka et al. (2007). In the meantime, several other studies on the validity of the OPD in terms of relationship diagnostics (Stasch, Schmal, Hillenbrand, & Cierpka, 2007), conflict diagnostics (Schneider, Mendler, Heuft, & Burgmer, 2008), and structure diagnostics (Böker et al., 2007) have been published. Overall, these studies show that the three psychodynamic axes of the OPD can be used to capture central psychoanalytic concepts and to describe mental dysfunctions.

OPD ratings are based on clinical interviews conducted in accordance with the OPD interview manual (OPD Task Force, 2001). Interviews were conducted by study collaborators (i.e., not by the therapists) and video recorded. The interview, which takes between 1 and 1.5 hr, focuses on the subjective

experiences and behaviour of the patient within personal relationships with the aim of identifying indications of patients' latent conflicts and structural features. For the purpose of assessing changes, the interviews were reconducted at the end of treatment and once again video recorded.

Based on the OPD assessment, core problem areas were defined for each patient. The term "core" refers to those aspects of the patient's OPD profile that were presumed to produce or sustain psychic and somatic symptoms and interpersonal difficulties. The definition of such problem areas thus carried the status of a psychodynamic hypothesis, inferring that change had to take place within these areas in order to effectuate a substantial reduction or disappearance of patients' symptoms and complaints. Because therapists were ignorant to this definition, they were consequently not able to gear their interventions toward alleviating problems in the patient's defined core areas. The items listed in Table I were used to determine these problem areas. Taking these items together (one core dysfunctional relationship pattern, eight types of conflict, and six dimensions of structural capacity/vulnerability) results in a total of 15 problematic features from which problem areas can be selected for each patient. Previous studies (Grande, Rudolf, Oberbracht, & Jakobsen, 2001; Grande et al., 2003) have indicated that a selection of five items from the overall OPD profile is sufficient to identify a patient's problem constellation. The selection of problem areas was carried out by two independent raters. Given the importance of valid item selection for this procedure, discrepancies between raters were discussed in order to reach a consensus. In the case of doubt, a third expert was included.

In a second step, the HSCS (Figure 1) was used to identify the way in which patients dealt with the selected problem areas. Ratings were performed based on the video-recorded OPD interviews. The HSCS is a modified form of the Assimilation of Problematic Experiences Scale (Stiles et al., 1990; Stiles, Meshiot, Anderson, & Sloan, 1992; Stiles,

Shapiro, Harper, & Morrison, 1995), which is more strongly oriented toward a psychoanalytic model of process and change (Rudolf et al., 2000). Each stage of the scale marks a therapeutically significant step, beginning with increasing awareness of a previously unperceived problem area, extending to the therapeutic working through of associated aspects and experiences, and then to subsequent basic changes in both the patient's experience and specific external behaviours. The scale was applied at the onset and the end of treatment to assess the way in which the patient dealt with the problem areas represented in the selected OPD items. This was carried out for each of the five problem areas. Based on the assumption that these problems interact with one another and collectively produce the patient's difficulties, mean structural change scores were calculated by averaging HSCS ratings across the five problem areas. Structural changes were thus represented as shifts in mean HSCS scores along the scale. For calculation purposes, intermediate scale points (e.g., 3- and 3+; see Figure 1) were assigned corresponding scores (e.g., 2.7 and 3.3, respectively).

To date, the HSCS has been used in two studies (Grande, Rudolf, Oberbracht, & Jakobsen, 2001; Grande et al., 2003; Schneider, Schmitz-Moormann, Bär, Driesch, & Heuft, 2006). In a study on 12-week inpatient treatment ( $N=49$ ), Grande, Rudolf, Oberbracht, and Jakobsen were able to show that initial HSCS values significantly correlated with prognostically favourable patient characteristics; pre-post changes in HSCS further correlated highly with outcome evaluations performed by members of the therapeutic team. In a 6-month follow-up with the same patients (Grande et al., 2003), those patients who had achieved good results on the HSCS upon finishing therapy showed progressive changes in important life domains, changes that were not made by those with poorer HSCS results. In a mixed group of 173 inpatient courses of treatment, Schneider et al. (2006) demonstrated that HSCS changes significantly correlated with the attainment

Table I. Operationalized Psychodynamic Diagnostic Axes and List of Potential Problem Areas

Relationship	Life-determining conflicts	Structural capacities/ vulnerabilities
Individualized formulation of a core dysfunctional relationship pattern	<ol style="list-style-type: none"> <li>1. Dependence/autonomy conflict</li> <li>2. Submission/control conflict</li> <li>3. Care/autarchy conflict</li> <li>4. Self-value conflicts</li> <li>5. Guilt conflicts</li> <li>6. Oedipal-sexual conflicts</li> <li>7. Identity conflicts</li> <li>8. Deficient awareness of feelings and conflicts</li> </ol>	<ol style="list-style-type: none"> <li>1. Capacity for experience of self</li> <li>2. Capacity for self-control</li> <li>3. Capacity for defence</li> <li>4. Capacity for object-experience</li> <li>5. Capacity for communication</li> <li>6. Capacity for attachment</li> </ol>

Stages	Manual excerpt
1. <b>Problem area warded off</b>	The problem is entirely unconscious; associated experiences are evaded; problematic behaviour is ego-syntonic; the patient has "no problem" with the problem area
2. <b>Unwanted preoccupation with the problem</b>	Unpleasant feelings and thoughts in connection with the problem area can no longer be immediately rejected, but preoccupation with the problem is reluctant; external confrontations with the problem take place, but are rejected as disturbances; the patient does not realise that problems might be associated with his/her own person
3. <b>Vague awareness of the problem</b>	Patient notices/suspects the existence of a problem that is part of him/herself and cannot simply be rejected; recurrence causes the problem to take on a continuing existence; negative affects originate from the tension between the insistent nature of the problem and the pat.'s defensive/aversive attitude
4. <b>Acceptance and exploration of the problem</b>	The problem begins to take on a new shape within the pat.'s consciousness; there are incipient indications of an active, "head-on" preoccupation with the problem; the problem can now be formulated as an "assignment" and can hence be made the subject of therapeutic work; destructive, rejective responses may interfere with this attitude but can no longer undermine it altogether
5. <b>Deconstruction in the problem area</b>	Querying and disintegration of accustomed coping modes; uncertainty concerning evaluations of own person and others; perception of own limitations and deficiencies; resignation and moods of despair alternate with urges toward reparation; old modes are lost and cut off, new ones not yet accessible
6. <b>Reorganization in the problem area</b>	Abandonment and final relinquishing of accustomed coping modes; pat. is increasingly self-reliant in his/her own experience and able to take control of and assume responsibility for his/her own life in the problem area; increasingly conciliatory approach to problem area; solutions materialize spontaneously and unexpectedly; re-integration
7. <b>Integration of the problem</b>	Dealing with the problem has become something natural; the area has lost its special significance in the eyes of the pat.; the problem is something which belongs to the past, preoccupies pat. as a memory

Figure 1. Heidelberg Structural Change Scale.



of individual therapy goals that had been defined by the patients themselves; they did not, however, correlate with pre–post changes on the Brief Symptom Inventory developed by Derogatis or the IIP total score.

Interrater reliability for the selection of the five OPD core problem areas and the HSCS rating was measured based on a total of six raters, all of whom were authors of the present paper (R. D., B. K., M. L., C. O., S. S., M. S.). In line with specifications of the OPD Task Force (2001), all raters received at least 60 hr of training in applying the OPD; performance checks were conducted based on pre-defined criteria (Grande, 2004) in addition to using four to six standard videos before raters began their work. Raters were subsequently trained to use the HSCS by two of the developers of the scale (T.G., C.O.). Compared with the OPD ratings, the HSCS represents a less demanding instrument and its application is rapidly learned. Performance checks were once again conducted using four to six standard videos, after which raters entered the study. Because the study extended over a period of several years, the entire group regularly (approximately four times per year) met to perform collective ratings and discuss discrepancies.

Each case was assessed by two raters, and ratings at the onset and end of therapy were performed by different groups of raters. In the case of material collected at the end of therapy, the Berlin group rated the Heidelberg material and visa versa. Raters assessing the HSCS at the end of treatment were informed about the problem areas that had been defined for each patient at the onset of treatment but otherwise had no baseline information about the patient or information concerning developments over the course of therapy. For the selection of problem areas from the OPD profile at the start of treatment, a kappa of .62 was calculated (Cohen, 1977). With regard to HSCS ratings, an interrater agreement of intraclass correlation coefficient (ICC) (1.1) = 0.83 was found.

The SCL-90-R (Derogatis, 1994; German version: Franke, 2002) was used by patients for self-assessment of symptom status. The GSI served as a measure of present overall distress. For the assessment of relationship problems, the German version of the IIP-Depression (IIP-D; Horowitz et al., 2000) was applied. Here, the total score was also used. Because of the heterogeneity of the patient group, it would not have been possible to test hypotheses on specific symptom clusters or specific interpersonal problems. Hence, therapy effects were not assessed with reference to SCL-90-R or IIP-D subscales. Both the SCL-90-R and the IIP-D were completed by patients at the start and end of therapy.

At the end of treatment as well as at 1-year and 3-year follow-up, patients further performed a retrospective evaluation of therapy outcome using an eight-item questionnaire covering various aspects of therapeutic change: mental symptoms, somatic symptoms, interpersonal problems, coping with life demands, overall capacity, enjoyment potential, self-esteem, and general contentment with life. Questionnaire instructions were as follows: “Please call to mind the troubles and difficulties from which you suffered upon beginning psychotherapy. When you view yourself now: To what extent have these troubles and difficulties changed since back then?” These instructions were followed by the just-listed aspects of change (e.g., “interpersonal problems”), which were to be rated on a 6-point scale with the following anchor points:  $-1$ , *deteriorated*;  $0$ , *unchanged*;  $+1$ , *slightly improved*;  $+2$ , *clearly improved*;  $+3$ , *considerably improved*;  $+4$ , *maximally improved*. Internal consistency (Cronbach’s alpha) for the total scale amounted to .94 at the end of treatment, .95 at 1-year follow-up, and .94 at 3-year follow-up. These values are high and justify the use of a summary score. Items were summated and divided by the number of items; resulting values ranged between  $-1$  and  $+4$ .

This instrument for retrospective outcome evaluation was developed in the context of a previous investigation on 162 courses of psychodynamic therapy (so-called Berlin Psychotherapy Study; cf. Rudolf, 1991). Patients in this study were requested to describe changes that in their eyes had been central in the form of free text. A qualitative analysis of these texts led to the identification of the change categories, which are assessed by the eight items included in the present instrument. Similar to other studies discussed early in this article, investigations using this instrument (Grande et al., 2001; Rudolf, Laszig, & Henningsen, 1997) have revealed close associations with other retrospective patient outcome evaluations, such as general assessments of success, but only weak relationships with prospectively measured pre–post changes and outcome assessments from other sources (the therapist; cf. Hill & Lambert, 2004; Lambert & Hill, 1994). The reason why retrospective change evaluations in our view nonetheless represent an important outcome criterion and a well-grounded option for the present study was discussed early in the article.

Mean values and standard deviations for the retrospective assessment scale were as follows:  $M = 2.00$  ( $SD = 1.01$ ) at the end of treatment,  $M = 1.94$  ( $SD = 1.01$ ) at 1-year follow-up, and  $M = 2.02$  ( $SD = 1.00$ ) at 3-year follow-up. On average, patients thus rated their situation as having “clearly improved” across all assessment occasions. In each

case, average ratings were spread approximately 1 scale point around the mean.

### Data Analysis

The investigation examined how well various pre-post changes were able to predict the retrospective outcome evaluations of patients at the end of treatment and at 1-year and 3-year follow-up. Pre-post changes were measured using the SCL-90-R (GSI), IIP-D (total score), and the HSCS (mean rating with respect to the five problems areas). Residual scores resulting from a regression of initial scores on scores at the end of therapy were used as a measure of change. This approach corrects for regression to the mean and is recommended for application when using individualized measures of change such as the HSCS (Beutler & Hamblin, 1986).

For the prediction of outcome evaluations, two series of hierarchical regression analyses were performed. In the first of these, the three residual change scores (GSIres, IIPtotalres, and HSCSres) were separately entered as predictors of the three outcome evaluations. The first predictor to be entered into each regression analysis was the measure of change (e.g., GSIres), followed by the interaction between the measure of change and the treatment (e.g., Treatment  $\times$  GSIres) for the purpose of examining the potential influence of treatment form on the prediction of outcome. With three measures of change and three outcome criteria, a total of nine analyses were necessary.

In the second series of regression analyses, the residual scores GSIres, IIPtotalres, and HSCSres were hierarchically entered into the analysis. To examine the potential influence of treatment form, interactions between treatments and the three measures of change (Treatment  $\times$  GSIres, Treatment  $\times$  IIPtotalres, and Treatment  $\times$  HSCSres) were entered in a final step. This procedure was consecutively conducted for the prediction of outcome evaluations at the end of treatment, at 1-year follow-up, and at 3-year follow-up. This series, therefore, comprised a total of three regression analyses.

We hypothesized that, in line with the asserted long-term effects of structural change, pre-post changes in the HSCS (HSCSres) would be significantly more predictive of outcome at 1-year and certainly at 3-year follow-up compared with pre-post changes in both IIPtotalres and GSIres. Predictive superiority of the HSCSres exclusively at 1-year follow-up (but not at 3-year follow-up) would thus not confirm our hypothesis. Predictive superiority of HSCSres exclusively at 3-year follow-up would limit

the validity of our hypothesis to very long-term effects. This would correspond with the expectation found in other psychoanalytic follow-up studies that long-term effects can only be examined 2 to 5 years after therapy completion (cf. Wallerstein, 2001). We, therefore, examined two hypotheses, a limited hypothesis (relating to the 3-year follow-up) and an extended hypothesis (additionally relating to the 1-year follow-up). Finally, differences between treatment forms were not expected to have any influence on the described associations, and the interaction between treatment and measures of change was not expected to be significant.

No assumptions were made regarding the relative strength of the associations between the three measures of change and the outcome criterion at the end of therapy; a comparison of these associations was, however, also calculated.

In testing the hypotheses, correlations between outcome criterion and HSCSres were compared with correlations between outcome criterion and IIPtotalres as well as with correlations between outcome criterion and GSIres at each of the three assessment occasions. The significance of observed differences ( $\Delta r_{ab} - r_{ac}$ ) was subsequently examined. Because two comparisons were required for testing each hypothesis, a  $p = .025$  level of significance was selected based on the Bonferroni correction. The correlations required for comparisons were calculated in Step 1 of each of the analyses in the first series of hierarchical regression analyses described previously (see also Table II).

In the second series of regression analyses (see Table III), the residual scores GSIres, IIP totalres, and HSCSres were hierarchically entered into the analysis. This analysis examines whether the HSCSres explains additional variance in the outcome criterion when the remaining predictors (IIPtotalres, GSIres) are accounted for. Again, it was assumed that the additional share of variance explained by HSCSres would certainly be significant at 3-year follow-up (limited hypothesis) and in an extended assumption also at 1-year follow-up (extended hypothesis).

### Results

Table II presents the results of the nine analyses from the first regression series, in which the predictive value of the three measures of change was separately investigated. For the prediction of outcome evaluations, pre-post changes in GSI accounted for 26.5% of criterion variance at the end of therapy, 9.7% at 1-year follow-up, and 1.4% at 3-year follow-up. The variable IIPtotalres accounted for 45.3% of variance at the end of therapy, 17.2% at 1-year follow-up, and

Table II. Summary of Nine Hierarchical Multiple Regression Analyses for the Prediction of Patient Evaluations of Treatment Outcome: Separate Analyses for Each Measure of Pre-Post Change

	Retrospective assessments of change		
	End of therapy (n = 54)	One-year follow-up (n = 53)	Three-year follow-up (n = 50)
Step 1:	$R = .51, R^2 = .26$	$R = .31, R^2 = .10$	$R = .12, R^2 = .01$
GSI <sub>res</sub>	$\Delta R^2 = .26, F(1, 52) = 18.70^{***}$	$\Delta R^2 = .10, F(1, 52) = 5.50^*$	$\Delta R^2 = .01, F(1, 52) = 0.66$
Step 2:	$R = .56, R^2 = .31$	$R = .31, R^2 = .10$	$R = .15, R^2 = .02$
GSI <sub>res</sub> x Treatm.	$\Delta R^2 = .04, \Delta F(1, 51) = 3.3$	$\Delta R^2 = .00, \Delta F(1, 51) = 0.01.$	$\Delta R^2 = .01, \Delta F(1, 51) = 0.47$
Step 1:	$R = .67, R^2 = .45$	$R = .41, R^2 = .17$	$R = .24, R^2 = .06$
IIP <sub>totalres</sub>	$\Delta R^2 = .45, F(1, 52) = 43.14^{***}$	$\Delta R^2 = .17, F(1, 52) = 10.60^{**}$	$\Delta R^2 = .06, F(1, 52) = 2.87$
Step 2:	$R = .67, R^2 = .45$	$R = .41, R^2 = .17$	$R = .34, R^2 = .11$
IIP <sub>totalres</sub> x Treatm.	$\Delta R^2 = .00, \Delta F(1, 51) = 0.10$	$\Delta R^2 = .00, \Delta F(1, 51) = 0.00$	$\Delta R^2 = .06, \Delta F(1, 51) = 2.99$
Step 1:	$R = .47, R^2 = .22$	$R = .50, R^2 = .25$	$R = .45, R^2 = .20$
HSCS <sub>res</sub>	$\Delta R^2 = .22, F(1, 52) = 14.44^{***}$	$\Delta R^2 = .25, F(1, 52) = 17.26^{***}$	$\Delta R^2 = .20, F(1, 52) = 11.90^{***}$
Step 2:	$R = .47, R^2 = .22$	$R = .52, R^2 = .27$	$R = .45, R^2 = .20$
HSCS <sub>res</sub> x Treatm.	$\Delta R^2 = .00, \Delta F(1, 51) = 0.00$	$\Delta R^2 = .02, \Delta F(1, 51) = 1.37$	$\Delta R^2 = .00, \Delta F(1, 51) = 0.04.$

Note. GSI<sub>res</sub> = residual scores of the Global Severity Score (SCL-90); IIP<sub>totalres</sub> = residual scores of the total score of the Inventory of Interpersonal Problems (IIP); residual scores of the Heidelberg Structural Change Scale (HSCS);  $R$  = multiple correlations;  $R^2$  = total explained variance;  $\Delta R^2$  = increase in explained variance associated with the step.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

5.6% at 3-year follow-up. In contrast, HSCS<sub>res</sub> accounted for a highly significant amount of variance in the prediction of all three outcome criteria: 21.7% at the end of therapy, 25.3% at 1-year follow-up, and 19.9% at 3-year follow-up. Interactions between treatment form and the three measures of change did not account for a significant amount of variance at any of the three retrospective assessment occasions. These results are also presented in Figure 2.

Differences between HSCS<sub>res</sub>-criterion correlations and IIP<sub>totalres</sub>-criterion correlations were as follows:  $z = -1.75$  ( $p = .042$ ) at the end of therapy,  $z = 0.64$  ( $p = .261$ ) at 1-year follow-up, and  $z = 1.40$  ( $p = .081$ ) at 3-year follow-up. The  $z$  value at the end

of therapy has a negative sign, thus indicating that the correlation between IIP<sub>totalres</sub> and the criterion was higher at this occasion than the correlation between HSCS<sub>res</sub> and the criterion (this effect was reversed at both follow-ups).

Differences between HSCS<sub>res</sub>-criterion correlations and GSI<sub>res</sub>-criterion correlations were as follows:  $z = -0.34$  ( $p = .367$ ) at the end of therapy,  $z = 1.30$  ( $p = .097$ ) at 1-year follow-up, and  $z = 2.04$  ( $p = .021$ ) at 3-year follow-up. The  $z$  value at the end of therapy once again has a negative sign, thus indicating that the correlation between GSI<sub>res</sub> and the criterion was higher at this occasion than the correlation between HSCS<sub>res</sub> and the criterion (this effect was again reversed at both follow-ups).

Table III. Summary of Three Hierarchical Multiple Regression Analyses for the Prediction of Retrospective Patient Evaluations of Treatment Outcome: Combined Analyses Including All Measures of Pre-Post Change

	Retrospective Assessments of change		
	End of therapy (n = 54)	One-year follow-up (n = 53)	Three-year follow-up (n = 50)
Step 1:	$R = .51, R^2 = .26$	$R = .31, R^2 = .10$	$R = .12, R^2 = 0.01$
GSI <sub>res</sub>	$\Delta R^2 = .26, \Delta F(1, 52) = 18.70^{***}$	$\Delta R^2 = .10, \Delta F(1, 51) = 5.50^*$	$\Delta R^2 = .01, \Delta F(1, 48) = 0.66$
Step 2:	$R = .68, R^2 = .46$	$R = .42, R^2 = .17$	$R = .24, R^2 = 0.06$
IIP <sub>totalres</sub>	$\Delta R^2 = .20, \Delta F(1, 51) = 18.89^{***}$	$\Delta R^2 = .08, \Delta F(1, 50) = 4.67^*$	$\Delta R^2 = .05, \Delta F(1, 47) = 2.29$
Step 3:	$R = .72, R^2 = .52$	$R = .56, R^2 = .32$	$R = .46, R^2 = 0.21$
HSCS <sub>res</sub>	$\Delta R^2 = .06, \Delta F(1, 50) = 6.25^*$	$\Delta R^2 = .14, \Delta F(1, 49) = 10.24^{**}$	$\Delta R^2 = .15, \Delta F(1, 46) = 8.88^{**}$
Step 4:			
GSI <sub>res</sub> x Treatment	$R = .73, R^2 = .54$	$R = .60, R^2 = .36$	$R = .51, R^2 = 0.26$
IIP <sub>totalres</sub> x Treatment	$\Delta R^2 = .01, \Delta F(3, 47) = 0.47$	$\Delta R^2 = .04, \Delta F(3, 46) = 0.99$	$\Delta R^2 = .05, \Delta F(3, 43) = 0.97$
HSCS <sub>res</sub> x Treatment			

Note. GSI<sub>res</sub> = residual scores of the Global Severity Score (SCL-90); IIP<sub>totalres</sub> = residual scores of the total score of the Inventory of Interpersonal Problems (IIP); residual scores of the Heidelberg Structural Change Scale (HSCS);  $R$  = multiple correlations;  $R^2$  = total explained variance;  $\Delta R^2$  = increase in explained variance associated with the step.

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

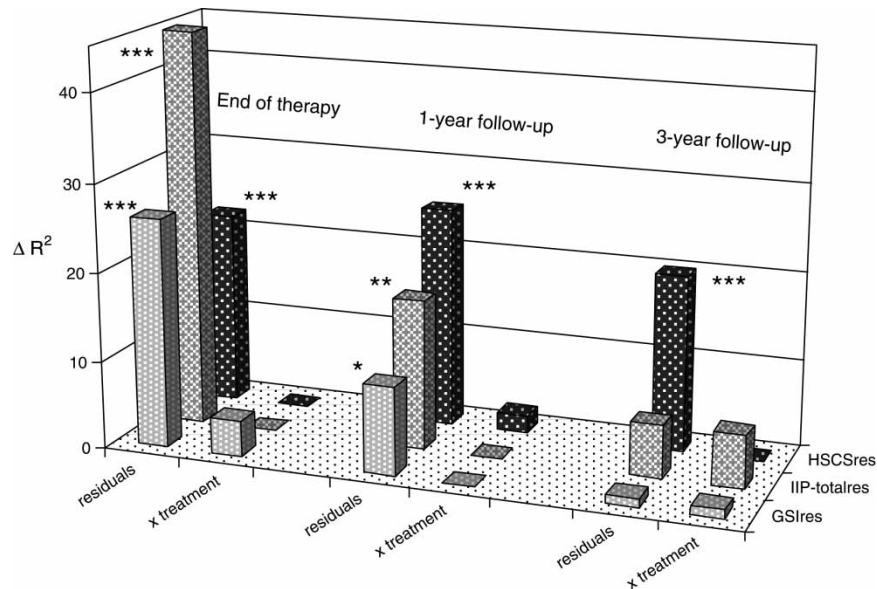


Figure 2. Results of 9 ( $3 \times 3$ ) hierarchical multiple regression analyses for the prediction of patient retrospective evaluations of treatment outcome. Separate analyses are conducted for each of the three residual change scores that were entered first to regression analysis, followed by the interaction between residual change scores and treatment. GSIres = residual scores of the Global Severity Score (SCL-90-R); IIPtotalres = residual scores of the total score of the Inventory of Interpersonal Problems (IIP); HSCSres = residual scores of the Heidelberg Structural Change Scale (HSCS);  $\Delta R^2$  = increase in explained variance. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

Table III presents regression results for analyses in which pre-post measures of change and their interactions with treatment form were hierarchically entered. The variable HSCSres again accounted for a significant or highly significant amount of variance in all three retrospective outcome evaluations, whereas GSIres and IIPtotalres accounted for large amounts of criterion variance at the end of therapy, considerably less at 1-year follow-up, and nonsignificant amounts at 3-year follow-up.

In light of the striking decrease in IIPtotalres-criterion and GSIres-criterion correlations from the end of therapy across 1-year follow-up and to 3-year follow-up, we finally examined post hoc whether these changes were statistically significant. Analyses were exclusively based on the 50 cases for which data were fully available at both assessment occasions. The difference between IIPtotalres-criterion correlations at the end of therapy and at 3-year follow-up was  $z = 3.43$  ( $p < .001$ ); between GSIres-criterion correlations at the end of therapy and at 3-year follow-up  $z = 2.84$  ( $p = .002$ ); and between HSCSres-criterion correlations at the end of therapy and at 3-year follow-up  $z = 0.42$  ( $p = .33$ ). There was thus no significant change in HSCSres-criterion correlations from the end of therapy to 3-year follow-up.

## Discussion

Results confirm that, in comparison with symptomatic changes, structural changes achieved by the

end of therapy are a significantly better predictor of retrospective outcome evaluations conducted by the patient 3 years after therapy completion. In other words, when patients are requested to evaluate their therapy-related progress in various life domains 3 years after finishing therapy, their evaluations are significantly better explained by the structural changes compared with the symptomatic changes achieved by the end of therapy. This effect was not found at 1-year follow-up.

Our hypotheses concerning the comparison between structural and interpersonal changes were not confirmed for either 3-year or 1-year follow-up. A post hoc analysis revealed, however, that, although the association between interpersonal changes and retrospective outcome evaluation was significantly stronger than that between structural changes and the criterion at the end of therapy, it clearly decreased and became even weaker than this relationship at 3-year follow-up. This decrease was highly significant and shows that when patients are asked to retrospectively evaluate the progress they have made as a result of therapy, their evaluations at the end of therapy are significantly more greatly determined by interpersonal changes than is the case 3 years later.

Pre-post structural change is the only measure that continues to (highly) significantly correlate with the patient's retrospective outcome evaluation after the end of treatment. In all analyses (i.e., at the end of therapy and at 1-year and 3-year follow-up) simultaneously including the three measures of

change as predictors, structural change significantly explained a unique amount of criterion variance.

We consider these results to be consistent with the psychoanalytic view that treatment should strive to achieve fundamental changes in the patient's personality; changes that have long-term effects in the life of the patient. Because change processes often continue beyond the end of therapy, the complete scope of these effects is not necessarily immediately visible to the patient. As a result, it is frequently only with hindsight that patients recognize and appreciate the value of their therapy. The patient's experience at the end of therapy is more strongly influenced by other therapeutic effects, namely by changes in the areas of symptom distress and relationship problems. When it comes to evaluating those specific therapeutic changes that will have a long-term effect on the patient's life at this point in time, a clinical expert assessing the structural changes seems able to provide a more reliable judgment, with which the patient somewhat "belatedly" then agrees.

In selecting retrospective patient outcome evaluations as outcome criterion in the present study, we chose to assign this measure a central role. As discussed early in this article, such measures have been challenged in the literature and judged rather critically in comparison to pre-post measures (Hill & Lambert, 2004; Lambert & Hill, 1994). In line with Michalak et al. (2003) and Schulte (2008), we argued that retrospective outcome evaluations capture something different from that which is reflected by prospectively measured pre-post changes, given that they are based on relative assessments that relate to individual goals rather than absolute assessments of therapeutic achievements. We further pointed out that individual goals and, therefore, the internal evaluation norms of the patients change over the course of therapy (cf. also Beutler & Hamblin, 1986) and that such changes to the inner reference system represent a particular aim of psychoanalytic treatment. We emphasize this point with reference to the reports of the patients at the follow-ups: Similar to other psychoanalytic follow-up studies (Leuzinger-Bohleber, 2002; Pfeffer, 1959), our patients repeatedly retrospectively reported that their views and evaluations had been fundamentally transformed during but also following their therapy and that they now saw themselves and other people "in a different light" so to speak. Problems that had caused much distress before therapy appeared less significant in hindsight and behaviours and arrangements that had previously seemed "normal" were perceived as problematic. That which Menninger (1958) wrote about the effects of psychoanalytic therapy based on the example of a fictitious case

would seem to apply to more than just a few of these patients: "Although it is true that his expectations were not met, his gains were beyond his expectations" (as cited in Wallerstein, 1965, S. 751). In our opinion, these complex processes of change can at least be globally assessed using retrospective evaluations, all the more so considering that no other options are available. At the same time, we are all too aware that retrospective evaluations of change show certain weaknesses. They demand that patients perform a complex cognitive operation that may impair the accuracy of their evaluation: They must first recall their level of distress at the onset of therapy (which may differ from what they reported at that time), assess their current level of distress, and subsequently determine the difference between the two. For future studies of this kind, it would in our view, therefore, seem desirable to use methods and strategies that allow for changing success criteria without having to accept the limitations of retrospective evaluations.

Some further methodological limitations must be noted: A number of therapists (14 of 45) contributed two courses of therapy to the study. Our data are, therefore, nested, a fact that was not accounted for in our statistical analyses. Diagnoses were not assessed in a standardized manner and their validity is, therefore, questionable. Moreover, because of the naturalistic design of the study, it was not possible to examine the way in which therapists actually worked within courses of therapies independently of information provided by the therapists themselves. Therapy duration and the number of sessions varied enormously in both treatments. Our study is, therefore, not able to answer the question concerning the therapeutic interventions and strategies needed to achieve structural change and the long-term benefits investigated: Therapeutic approach had no influence; interaction effects between the two forms of psychodynamic treatment and pre-post changes were not observed in any of our analyses.

In our view, the approach adopted in the present study enables a comparison of different types of pre-post changes with regard to short- and long-term therapeutic effects. The results of such comparisons carry consequences for practical clinical work: If certain pre-post changes are associated with more long-term effects, then therapy should aim to work toward effectuating these very changes. The question concerning what the therapist can do to contribute to such changes thus represents an important issue to be addressed in future investigations.

## References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Beutler, L. E., & Hamblin, D. L. (1986). Individualized outcome measures of internal change: Methodological considerations. *Journal of Consulting and Clinical Psychology, 54*, 48–53.
- Böker, H., Schopper, C., Straub, M., Himmighoffen, H., Endrass, J., Küchenhoff, B., et al. (2007). Automutilistisches Verhalten bei Patientinnen mit affektiven Störungen: Untersuchung der Persönlichkeitsstruktur und Affektregulation mittels Operationalisierter Psychodynamischer Diagnostik (OPD) [Automutilistic behaviour in female patients with affective disorders: investigation of personality structure and affect regulation by means of Operationalized Psychodynamic Diagnostics (OPD)]. *Psychotherapie. Psychosomatik und Medizinische Psychologie, 57*, 319–327.
- Brockmann, J., Schlüter, T., Brodbeck, D., & Eckert, J. (2002). Die Effekte psychoanalytisch orientierter und verhaltenstherapeutischer Langzeittherapien [Effects of psychoanalytically oriented and of behavioural long-term therapies. A comparative study]. *Psychotherapeut, 47*, 347–355.
- Cierpka, M., Grande, T., Rudolf, G., van der Tann, M., Stasch, M., & OPD Task Force. (2007). The Operationalized Psychodynamic Diagnostics system: Clinical relevance, reliability and validity. *Psychopathology, 40*, 209–220.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Connolly Gibbons, M. B., Crits-Christoph, P., Barber, J. P., & Schamberger, M. (2007). Insight in psychotherapy: A review of empirical literature. In L. G. Castonguay & C. Hill (Eds), *Insight in psychotherapy* (pp. 143–165). Washington, DC: American Psychological Association.
- Derogatis, L. (1994). *Symptom Checklist-90-Revised: Administration, scoring, and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.
- DeWitt, K. N., Hartley, D., Rosenberg, S. E., Zilberg, N. J., & Wallerstein, R. S. (1991). Scales of Psychological Capacities: Development of an assessment approach. *Psychoanalytic & Contemporary Thought, 14*, 343–361.
- Dilling, H., Mombour, W., Schmidt, M. H., & Schulte-Markwort, E. (Eds). (1994). *Internationale Klassifikation psychischer Störungen ICD-10, Kapitel V(F). Forschungskriterien* [International Classification of Mental Disorders ICD-10, Chapter V(F). Research criteria]. Bern, Germany: Huber.
- Fonagy, P., Target, M., Steele, M., & Steele, H. (1998). *Reflective-functioning manual: For application to Adult Attachment Interviews. Version 5.0*. London: University College London.
- Franke, G. (2002). *SCL-90-R: Die Symptom-Checkliste von Derogatis (2. Aufl.)* [The Derogatis Symptom Checklist (2nd ed.)]. Göttingen, Germany: Beltz.
- Grande, T. (2004). Zur Beurteilung der Leistung einzelner OPD-Rater und Standards für die Verwendung der OPD in der empirischen Forschung [Qualification of OPD raters and the use of the OPD in empirical research]. In R. W. Dahlbender, P. Buchheim, & G. Schüssler (Eds), *Lernen an der Praxis. OPD und Qualitätssicherung in der Psychodynamischen Psychotherapie* [Learning from practice. OPD and quality assurance in psychodynamic psychotherapy] (pp. 245–250). Bern, Germany: Huber.
- Grande, T., Dilg, R., Jakobsen, T., Keller, W., Krawietz, B., Langer, M., et al. (2006). Differential effects of two forms of psychoanalytic therapy—Results of the Heidelberg-Berlin Study. *Psychotherapy Research, 16*, 470–485.
- Grande, T., Rudolf, G., Oberbracht, C., & Jakobsen, T. (2001). Therapeutische Veränderungen jenseits der Symptomatik. Wirkungen stationärer Psychotherapie im Licht der Heidelberger Umstrukturierungsskala [Therapeutic changes beyond symptoms. The effects of inpatient psychotherapy in light of the Heidelberg Structural Change Scale]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie, 47*, 213–233.
- Grande, T., Rudolf, G., Oberbracht, C., Jakobsen, T., & Keller, W. (2004). Investigating structural change in the process and outcome of psychoanalytic treatment. The Heidelberg-Berlin Study. In P. Richardson, H. Kächele, & C. Renlund (Eds), *Research on psychoanalytic psychotherapy with adults* (pp. 35–61). London: Karnac.
- Grande, T., Rudolf, G., Oberbracht, C., & Pauli-Magnus, C. (2003). Progressive changes in patients' lives after psychotherapy: Which treatment effects support them? *Psychotherapy Research, 13*, 43–58.
- Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcome. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (pp. 84–135). New York: Wiley.
- Horowitz, L. M., Strauss, B., & Kordy, H. (2000). *Inventar zu Erfassung interpersonaler Probleme (2. Aufl.)* [Inventory of Interpersonal Problems, 2nd ed.]. Göttingen, Germany: Beltz.
- Huber, D., Henrich, G., & Klug, G. (2007). The Inventory of Interpersonal Problems (IIP): Sensitivity to change. *Psychotherapy Research, 17*, 474–481.
- Kendall, P. C., Holmbeck, G., & Verduin, T. (2004). Methodology, design, and evaluation in psychotherapy research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (pp. 16–43). New York: Wiley.
- Kernberg, O. F. (1991). Psychic structure and structural change: An ego psychology-object relations theory viewpoint. In T. Shapiro (Ed.), *The concept of structure* (pp. 315–337). Madison, CT: International Universities Press.
- Kernberg, O. F. (1999). Psychoanalysis, psychoanalytic psychotherapy, and supportive psychotherapy: Contemporary controversies. *International Journal of Psychoanalysis, 80*, 1075–1091.
- Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin & S. L. Garfield (Eds), *Handbook of psychotherapy and behavior change* (pp. 72–113). New York: Wiley.
- Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behaviour change* (pp. 139–193). New York: Wiley.
- Leuzinger-Bohleber, M. (2002). A follow-up study—Critical inspiration for our clinical practice? In M. Leuzinger-Bohleber & M. Target (Eds), *Outcomes of psychoanalytic treatment* (pp. 143–173). London: Whurr.
- Leuzinger-Bohleber, M., Stuhr, U., Rüger, B., & Beutel, M. (2001). Langzeitwirkungen von Psychoanalysen und Psychotherapien. Eine multiperspektivische, repräsentative Katamnesestudie [Long-term effects of psychoanalysis and psychotherapy. A multi-center representative follow-up study]. *Psyche, 55*, 193–276.
- Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Clinical and Consulting Psychology, 66*, 400–410.
- Lunnen, K. M., Ogles, B. M., & Pappas, L. N. (2008). A multiperspective comparison of satisfaction, symptomatic change, perceived change, and end-point functioning. *Professional Psychology: Research and Practice, 39*, 145–152.
- Menninger, K. A. (1958). *Theory of psychoanalytic technique*. New York: Basic Books.
- Michalak, J., Kosfelder, J., Meyer, F., & Schulte, D. (2003). Messung des Therapieerfolgs. Veränderungsmessung oder retrospektive Erfolgsbeurteilung [Measuring the success of

- therapy. Change measures or retrospective assessment]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 32, 94–103.
- Moore, B. E., & Fine, B. D. (1990). *Psychoanalytic terms & concepts*. New Haven, CT: Yale University Press.
- Nedelmann, C. (1981). Behandlungsziel und Gesundheitsbegriff in der Psychoanalyse [Treatment goals and the concept of health in psychoanalysis]. In H. Bach (Ed.), *Der Krankheitsbegriff in der Psychoanalyse* [The concept of disease in psychoanalysis] (pp. 55–67). Göttingen, Germany: Vandenhoeck & Ruprecht.
- OPD Task Force. (Ed.). (2001). *Operationalized Psychodynamic Diagnostics. Foundations and manual*. Seattle, WA: Hogrefe & Huber.
- OPD Task Force. (Ed.). (2008). *Operationalized psychodynamic diagnosis (OPD-2). Manual of diagnosis and treatment planning*. Cambridge, MA: Hogrefe & Huber.
- Pekarik, G., & Wolff, C. B. (1996). Relationship of satisfaction to symptom change, follow-up adjustment, and clinical significance. *Professional Psychology: Research and Practice*, 27, 202–208.
- Pfeffer, A. Z. (1959). A procedure for evaluating the results of psychoanalysis: A preliminary report. *Journal of the American Psychoanalytic Association*, 7, 418–444.
- Rapaport, D. (1960). *The structure of psychoanalytic theory: A systemizing attempt* (Psychological Issues Series, Vol. II). New York: International Universities Press.
- Rudolf, G. (1991). *Die therapeutische Arbeitsbeziehung* [The psychotherapeutic working alliance]. Berlin: Springer.
- Rudolf, G., Grande, T., Dilg, R., Jakobsen, T., Keller, W., Oberbracht, C., et al. (2002). Structural changes in psychoanalytic therapies—The Heidelberg-Berlin Study on Long-Term Psychoanalytic Therapies (PAL). In M. Leuzinger-Bohleber & M. Target (Eds.), *Longer-term psychoanalytic treatment—Perspectives for therapists and researchers* (pp. 201–222). London: Whurr.
- Rudolf, G., Grande, T., & Oberbracht, C. (2000). Die Heidelberger Umstrukturierungsskala. Ein Modell der Veränderung in psychoanalytischen Therapien und seine Operationalisierung in einer Schätzskaala [The Heidelberg Structural Change Scale. A model for change in psychoanalytic therapies and its operationalization in an assessment scale]. *Psychotherapeut*, 45, 237–246.
- Rudolf, G., Laszig, P., & Henningsen, C. (1997). Dokumentation im Dienste von klinischer Forschung und Qualitätssicherung [Documentation as a way of clinical research and quality control]. *Psychotherapeut*, 42, 145–155.
- Rüger, U., Dahm, A., & Kallinke, D. (2003). *Faber-Haarstrick: Kommentar Psychotherapierichtlinien (6. Aufl.)* [The Faber-Haarstrick Commentary on Psychotherapy Guidelines, 6th ed.]. Munich: Urban & Fischer Verlag.
- Sandler, J., & Sandler, A. (1983). The second censorship, the three box model, and some technical implications. *International Journal of Psychoanalysis*, 64, 413–425.
- Schauenburg, H., & Strack, M. (1999). Measuring psychotherapeutic change with the Symptom Checklist-90-R—SCL 90 R. *Psychotherapy and Psychosomatics*, 68, 199–207.
- Schlesinger, N., & Robbins, F. (1975). The psychoanalytic process: Recurrent patterns of conflicts and changes in ego functions. *Journal of the American Psychoanalytic Association*, 23, 761–782.
- Schneider, G., Mendler, T. M., Heuft, G., & Burgmer, M. (2008). Validität der Konfliktachse der Operationalisierten Psychodynamischen Diagnostik (OPD-1)—Empirische Ergebnisse und Folgerungen für die OPD-2 [Validity of axis III “conflicts” of Operationalized Psychodynamic Diagnostics (OPD-1)—Empirical results and conclusions for OPD-2]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 54, 46–62.
- Schneider, G., Schmitz-Moormann, S., Bär, O., Driesch, G., & Heuft, G. (2006). Psychodynamische Therapiefoki in einer stationären multimodalen Therapie—Zusammenhänge zu den Selbstauskunftsskalen und individuellen Therapiezielen [Psychodynamic therapy foci in inpatient psychotherapy—Relations to self-assessment scales and individual therapeutic goals]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 52, 259–273.
- Schulte, D. (2008). Patients’ outcome expectancies and their impression of suitability as predictors of treatment outcome. *Psychotherapy Research*, 18, 481–494.
- Stasch, M., Schmal, H., Hillenbrand, E., & Cierpka, M. (2007). Fokusorientierte Interventionen mit der OPD in der stationären Psychotherapie. Effekte auf Ergebnis und Verlauf der Behandlung [Focal interventions based on the OPD in inpatient treatment. Effects on outcome and process]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 53, 309–323.
- Stiles, W. B., Elliot, R., Llewelyn, S. P., Firth-Cozens, J. A., Margison, F. R., Shapiro, D. A., et al. (1990). Assimilation of problematic experiences by clients in psychotherapy. *Psychotherapy*, 27, 411–420.
- Stiles, W. B., Meshiot, C. M., Anderson, T. M., & Sloan, W. W. (1992). Assimilation of problematic experiences: The case of John Jones. *Psychotherapy Research*, 2, 81–101.
- Stiles, W. B., Shapiro, D. A., Harper, H., & Morrison, L. A. (1995). Therapist contributions to psychotherapeutic assimilation: An alternative to the drug metaphor. *British Journal of Medical Psychology*, 68, 1–13.
- Thomä, H., & Kächele, H. (1987). *Psychoanalytische Praxis: 1. Prinzipien*. Berlin: Springer.
- Ticho, G. R. (1967). On self-analysis. *International Journal of Psychoanalysis*, 48, 308–318.
- Wallerstein, R. S. (1965). The goals of psychoanalysis. *Journal of the American Psychoanalytic Association*, 13, 748–770.
- Wallerstein, R. S. (1988). Psychoanalysis and psychotherapy: Relative roles reconsidered. *Annual of Psychoanalysis*, 16, 129–151.
- Wallerstein, R. S. (1996). *Scales of Psychological Capacities. Version 1*. Unpublished manuscript.
- Wallerstein, R. S. (2001). Die Generationen der Psychotherapieforschung—ein Überblick [Psychoanalytic therapy research: An overview]. In U. Stühr, M. Leuzinger-Bohleber, & M. Beutel (Eds.), *Langzeitpsychotherapie. Perspektiven für Therapeuten und Wissenschaftler* [Long-term psychotherapy: Perspectives for therapists and researchers] (pp. 38–60). Stuttgart, Germany: Kohlhammer.
- Weinryb, R. M., & Rössel, R. J. (1991). Karolinska Psychodynamic Profile—KAPP. *Acta Psychiatrica Scandinavica*, 83, 1–23.