

## From symptom relief to interpersonal change: Treatment outcome and effectiveness in inpatient psychotherapy

MATTHIAS HAASE<sup>1</sup>, JÖRG FROMMER<sup>1</sup>, GABRIELE-HELGA FRANKE<sup>2</sup>, THILO HOFFMANN<sup>3</sup>, JÖRG SCHULZE-MUETZEL<sup>3</sup>, SUSANNE JÄGER<sup>4</sup>, HANS-JÖRGEN GRABE<sup>5</sup>, CARSTEN SPITZER<sup>5</sup>, & NORBERT SCHMITZ<sup>6</sup>

<sup>1</sup>*Psychosomatic Medicine and Psychotherapy, Otto-von-Guericke University, Magdeburg, Germany;* <sup>2</sup>*Department of Rehabilitation Psychology, University of Applied Science, Stendal, Germany;* <sup>3</sup>*Department of Psychosomatic Medicine and Psychotherapy, AWO Hospital, Jerichow, Germany;* <sup>4</sup>*Angewandte Humanwissenschaften, Hochschule Magdeburg-Stendal, Stendal, Germany* & <sup>5</sup>*Department of Psychiatry and Psychotherapy, Ernst-Moritz-Arndt University, Greifswald, Germany,* and <sup>6</sup>*Department of Psychiatry, McGill University, Montreal, Quebec, Canada*

(Received 21 September 2007; revised 21 March 2008; accepted 9 April 2008)

### Abstract

This study evaluated the impact of psychodynamic inpatient psychotherapy on patients' psychological distress and interpersonal problems during the course of treatment and 1 year later. A total of 156 patients were assessed with the Symptom Checklist-90-Revised and the Inventory of Interpersonal Problems at intake, 4 weeks later, and at the end of therapy. The follow-up assessment was conducted 1 year later. Results support psychodynamic approaches as well as the phase model, which stresses that the goals to be achieved by psychotherapeutic interventions are not only improvement of well-being and symptoms but also changes in interpersonal behavior. Consequently, on a long-term basis, the first 4 weeks of therapy seem to be insufficient, especially for adequate changes on the interpersonal level.

**Keywords:** anxiety; depression; dream work; eating disorders; psychoanalytic/psychodynamic therapy; trauma; psychotherapist training/supervision/development; process research

The effects of psychotherapeutic treatment are largely dependent on the number of treatment hours. On the basis of a meta-analysis of more than 2,431 therapies, Howard, Kopta, Krause, and Orlinsky (1989) hypothesized a logarithmic relationship between therapeutic dose and therapeutic effect, according to which the same dose of psychotherapy at the beginning of a treatment achieves essentially greater effects than in advanced phases.

Taking into account the general critiques of the drug metaphor implicit in the hypothesis of the dosage model (Stiles & Shapiro, 1989), Howard, Lueger, Maling, and Martinovich (1993) later developed the phase model, which states that effects in advanced phases are not necessarily lower but different compared with those in earlier phases. Interpreting the dosage model (Grissom, Lyons & Lutz, 2002) in this way, the phase model suggests that after a small number of psychotherapy sessions

(<8), a remoralization (improvement in the patient's subjective sense of well-being) can be attained. During a medium-length treatment (9–26 sessions), a remediation (improvement of symptoms) is seen, whereas the desired restoration at the level of social functioning—the learning of new roles (rehabilitation)—requires longer periods of time (>27 sessions). In terms of the rate of improvement, an average of eight sessions, according to these authors, is necessary to achieve a 50% rate of effectiveness, whereas a dose of 26 sessions would correspond to a 75% rate of effectiveness. Interestingly, the necessary number of treatment sessions in each case depends less on the therapeutic methods applied than on the chosen criteria of success (Lambert & Ogles, 2004). In addition, national differences and associated social, cultural, and economic factors may play a decisive role, as a German translator of a study by Lueger (1995) remarked.

Correspondence concerning this article should be addressed to Jörg Frommer, Psychosomatic Medicine and Psychotherapy, Otto-von-Guericke University, Leipziger Strasse 44, Magdeburg D-39120, Germany. E-mail: joerg.frommer@medizin.uni-magdeburg.de

For example, psychotherapeutic treatment in Germany differs from such treatment in the Anglo-American world by at least three features: (a) greater influence of psychodynamic therapeutic methods; (b) longer duration of treatment; and (c) wider availability of inpatient psychotherapy. The legitimacy of these special features is rooted in empirical findings that show that higher therapeutic doses also can be meaningful (i.e., that only 50% of patients who were limited to 21 therapeutic sessions evinced any clinically significant improvement; Lambert, Hansen, & Finch, 2001). Moreover, these studies have drawn attention to the fact that not only cognitive-behavioral therapy but also psychodynamic therapy in an outpatient setting is effective (Leichsenring, 2005) and that long-term therapy can also achieve specific effects (Bachrach, Galatzer-Levy, Skolnikoff, & Waldron, 1991). Furthermore, newer studies show that one third to three fourths of patients receiving hospital care profit from the inpatient setting (Bräutigam, Senf, & Kordy, 1990; Franz et al., 2000; Junge & Ahrens, 1996; Klose, Matteucci-Gothe, & Linden, 2006; Rudolf, Jakobsen, Micka, & Schumann, 2004).

However, research on inpatient psychotherapy encounters some methodological problems such as differences in treatment lengths (i.e., short vs. long term) and settings (rehabilitation hospital, psychiatric ward, psychosomatic ward) as well as variations in therapeutic intensity. Unfortunately, the majority of studies operationalize therapeutic intensity as duration of the inpatient treatment in days. One critical reservation against this type of definition is that it fails to take into account the varying degrees of temporal concentration of treatment in these hospitals and that dose-effect comparisons are not possible with outpatient care. Therefore, it might be helpful to measure the therapeutic dose not only in terms of days spent in the hospital but in terms of applied therapy hours.

In the conceptual scheme of our own study, we also focused on personality dimensions as reflected by learning of new roles and interpersonal behavior. We justify this decision by pointing out that efforts to influence role-conflict patterns anchored in patients' life histories and repeated in various interpersonal relationships call for a longer and more intensive psychotherapy than the mere improvement of symptoms. This is recognized not only in the phase model but also in psychodynamic thinking.

Furthermore, this is not simply a theoretical assumption, because in the meantime there is concrete empirical evidence that an amelioration of interpersonal problems, as measured by the Inventory of Interpersonal Problems (IIP-C; Horowitz, Strauß, & Kordy, 2000), requires higher doses of

therapy than needed to improve symptoms (Barkham, Rees, Stiles, Hardy, & Shapiro, 2002), as assessed by the Symptom Checklist-90-R (SCL-90-R; Derogatis, 1977). Because the phase model suggests different treatment stages, we decided not only to assess inpatients' psychopathology and interpersonal problems at admission and dismissal but to include assessments 4 weeks after the onset of treatment and 1 year after the termination of inpatient therapy as well.

Finally, for the reasons indicated previously, we defined the therapeutic dose in terms of applied therapy hours. The amelioration of interpersonal problems can also, of course, be used as an indicator for the lasting improvement of symptoms (Davies-Osterkamp, Strauss, & Schmitz, 1996). In a review of therapeutic failures, Mohr (1995) reports that interpersonal problems are often an indicator of negative therapeutic outcomes. Taking these considerations into account, our study focused on five hypotheses. First, most patients' symptoms considerably improve in the first 4 weeks of inpatient treatment. Second, between the 4-week assessment and the end of the inpatient stay, a general further improvement in symptoms is observed; however, this change is slighter than that noted in the first 4 weeks. Third, in terms of interpersonal problems, patients in therapy show statistically significant improvements, which remain less prominent than the overall improvement in symptoms. Fourth, improvements in interpersonal skills do not regress from the end of treatment to the date of the follow-up study 1 year later, at least according to individual scales of measurement of the IIP-C. Fifth, in regard to heretofore published data in the domain of interpersonal problems, we assume that patients showing no improvement in the area of interpersonal relations during therapy will likewise evince no lasting improvement of symptoms.

## Method

### Sample

The patients included in this study underwent treatment in the Acute Psychosomatic Ward at Jerichow Hospital (State of Saxony-Anhalt, Germany) from the first quarter of 2001 to the fourth quarter of 2003. Altogether, 408 men and women had been treated in this period. Sociodemographic and psychometric data were collected at the beginning of the inpatient therapy ( $t_0$ ), with data from the latter category again taken 4 weeks after the onset of therapy ( $t_1$ ) and at discharge ( $t_2$ ). Twelve months after the termination of the treatment,

follow-up questionnaires ( $t_3$ ) were mailed to all study participants.

A total of 156 patients were willing to take part in the follow-up assessment and returned the completed questionnaires (response rate: 38.2%). Failure to respond to the patient questionnaire could be attributed to various reasons: 84 patients (20.6%) could not be reached by mail, and 168 patients (41.2%) did not take part in the posttreatment follow-up because of a lack of interest, which they had already expressed at the end of the therapy.

The analysis of sociodemographic variance between the posttreatment follow-up group ( $N=156$ ) and the patient group not taking part in the posttreatment follow-up ( $M=41.5$ ,  $SD=12.3$ ) revealed that patients who did take part in the follow-up survey were statistically significantly older,  $F(1, 407)=5.03$ ,  $p<.025$ , than the nonparticipating patients. Moreover, the two groups of patients showed statistically significant differences in terms of the highest year of education completed,  $F(1, 407)=5.31$ ,  $p<.022$ . The greatest difference concerned the percentage of college/university graduates: In the posttreatment follow-up group, the figure was 19.9% and among the dropout group, 9.5%.

The final sample comprised 107 women (68.6%) and 49 men (31.4%). At the beginning of treatment, 50 (32.1%) patients were either single or in their first marriage, 19 (12.2%) were in their second marriage, and 26 (16.7%) were either separated or divorced. At the beginning of therapy, 120 (76.9%) patients were employed, 18 (11.5%) were unemployed, and 15 (9.6%) were on long-term sick leave. The average duration of inpatient stay was 60.5 days ( $SD=24.5$  days; range =8–128 days).

## Intervention

Psychodynamic individual therapy was carried out by experienced psychological or medical psychotherapists as well as by special therapists (music, art, and social therapy). When determining the therapeutic dose, we followed the consensus recommendations of the leading psychosomatic clinics with a psychodynamic approach. The consensual process is based on the planned political introduction of case mixed funding for inpatient treatment. According to this guideline, the treatment comprises medical therapy; psychodynamic individual therapy, which is goal related and temporally limited to 100 min/week; group psychotherapy at 120 min/week; and nonpsychotherapeutic, including music therapy, communicative movement therapy, art therapy, and various relaxation techniques (360 min/week; Heuft

et al., 2002). The average dose of therapy amounts to almost 10 hr/week.

## Instruments

Upon reception, a routine medical history was taken, followed by the collection of sociodemographic data. Psychological distress was recorded with the help of the revised self-report SCL-90-R comprising 90 items. The SCL-90-R was developed by Derogatis (1977), translated into German, and standardized on the basis of data representative of the national population (Franke, 2002). Its reliability in clinical groups is very high. Despite criticism pertaining to the factors of its nine subscales, which cannot always be factor analytically replicated (Hessel, Schumacher, Geyer, & Brähler, 2001), the SCL-90-R has attained the status of a standard instrument in psychotherapy research (Lang & Hoyer, 2003).

Interpersonal problems were assessed with the IIP-C (Horowitz et al., 2000). This procedure allows a differentiated diagnosis of interpersonal problems, particularly in the clinical field.

## Statistical Evaluation

The mean values of the scale established in the course of time were then simultaneously measured using multivariate analysis of variance (MANOVA). If any statistical significance stood out, a Bonferroni-corrected item-by-item individual comparison on all scales was done. The classification of effect size<sup>1</sup> was reached as follows: Mild effects are considered to be those up to .40, moderate effects are within the range of 0.40 to 0.80, and strong effects exceed .80 (Cohen, 1988).

Statistically significant changes were determined with the help of the Global Severity Index (GSI) of the SCL-90-R and the IIP Composite (IIP-C) score. To make sure that the findings reached were neither flaws in measurement nor accidents, the changed value had to surpass the reliable change index (RCI). For the normative group representative of the national population as well as for the 5.057 psychotherapeutic inpatients, the RCI for the GSI (SCL-90-R) is fixed at  $T\geq 5$  (Franke, 2002). Concerning the IIP-C, a clinical population of psychotherapeutic patients ( $N=303$ ) represents the normative group for the calculation of the RCI, which is fixed at a composite base value of 0.33 or higher (Wuchner, Eckert, & Biermann-Ratjen, 1993). A change was classified as a clinically significant improvement if the values from the onset of therapy, versus those at a later point of measurement ( $t_1$ ,  $t_2$ , or  $t_3$ ), showed a statistically significant improvement and if the patient was not within a

dysfunctional population group at the time of the later measurement (Schmitz & Davies-Osterkamp, 1997). On the SCL-90-R, the cutoff score stands at 0.70 in the case of psychotherapeutic inpatients; for the IIP-C, a cutoff score of 1.535 was calculated both for the normative group and for the clinical population. Consequently, in the field of symptoms (SCL-90-R), a change of measurement is classified as a clinically significant improvement if the test participants at the onset of therapy ranged higher than a  $T$  value of 60, and at a later point in time ( $t_1$ ,  $t_2$  or  $t_3$ ) below 60; in addition, the difference between both values must be  $T > 5$ . In the case of interpersonal problems (IIP-C), an improvement is considered clinically significant if a patient scores above a composite base value of 1.535 on the IIP-C scale at a given point in therapy ( $t_0$ ,  $t_1$ , or  $t_2$ ) and below this cutoff at a later assessment ( $t_1$ ,  $t_2$  or  $t_3$ ), while simultaneously having improved by a base value of at least .33.

## Results

### Diagnoses

The psychodiagnosis was reached with the aid of International Classification of Diseases (10th edition [ICD-10]) investigative criteria applied within a detailed anamnesis carried out by approved psychotherapeutic or medical personnel. As in the investigation by Rudolf et al. (2004), a strict separation was observed between symptomatology and personality structure, as postulated in psychodynamics. The diagnostic range of the present study comprises patients who were suffering from depressive disorders (ICD-10 F32-F34, 32%), acute stress-related disorders and adjustment disorders (ICD-10 F43, 16.7%), anxiety disorders (ICD-10 F40-F41, 20.5%), somatoform disorders (ICD-10 F45, 24.9%), eating disorder (ICD-10 F50, 4.5%), and "other" (1.4%). The three most common diagnoses were moderate depressive episodes (23.7%), somatization disorders (16.0%), and dissociative motor disorders (13.5%). One hundred seven (69%) patients taking part in the follow-up survey received a second ICD-10 diagnosis, and 26 (17%) received a third diagnosis.

### SCL-90-R Mean Values

The mean values of the SCL-90-R at the beginning of therapy revealed a marked symptomatic strain on all scales as well as among the global indexes. The scales for obsessive-compulsive traits ( $T = 65.62$ ,  $SD = 10.26$ ), depression ( $T = 67.83$ ,  $SD = 9.2$ ), and anxiety ( $T = 66.69$ ,  $SD = 9.74$ ), as well as the global indexes, were markedly high (Figure 1). Four weeks

after the beginning of therapy ( $t_1$ ), the improvements in the symptom strain were already visible. Only the GSI ( $T = 60.13$ ,  $SD = 10.22$ ) as well as the scales measuring depression ( $T = 60.74$ ,  $SD = 10.28$ ) and anxiety ( $T = 60.38$ ,  $SD = 10.36$ ) remained slightly raised. At the end of treatment ( $t_2$ ), a significant improvement in the symptomatic complaints had occurred. At the time of this measurement, all scales and the global indexes were in the normal range ( $T < 60$ ). These positive changes showed a slight decline by the posttreatment follow-up, but still remained well within the normal range.

As presented in Table I, the analyses of variance (ANOVAs) over the course of treatment revealed clear, statistically significant changes (MANCOVA with repetition of measurement on all nine scales of the SCL-90-R),  $F(3, 465) = 40.39$ ,  $p < .0001$ , 20.7% calculated variance resolution (VR). The extent of VR in the course of time was highest on the scale for depression, at 31.30% VR,  $F(3, 465) = 0.74$ ,  $p < .0001$ , and lowest on the scale for paranoid ideation, with a VR of 13.90%,  $F(3, 465) = 25.12$ ,  $p < .0001$ .

The Bonferroni-corrected paired individual comparisons applied successively on the subscales demonstrated that the patients underwent statistically significant changes, marked on all subscales and including the global indexes of the SCL-90-R, from the beginning of therapy ( $t_0$ ) to the measurements conducted 4 weeks after the onset of therapy ( $t_1$ ), the date of dismissal ( $t_2$ ), and the posttreatment follow-up ( $t_3$ ). Like the previous measurements, those undertaken between the 4-week assessment ( $t_1$ ) and at dismissal ( $t_2$ ) proved that a significant change had taken place.

### Effect Size

On the GSI ( $d_{GSI} = 0.67$ ), which is frequently used to measure the temporal course of the symptom strain, effects in the moderate range were recorded as well as for almost all other scales and global indexes in the period from admission to the measurements taken 4 weeks after onset of therapy. At discharge, high effects were found ( $d_{GSI} = 1.16$ ) followed by a slight decrease at the posttreatment follow-up ( $d_{GSI} = 0.80$ ).

### Clinical and Statistical Significance

Following criteria established by Jacobson and Truax (1991), 30.8% of patients were classified as clinically improved in their symptom strain 4 weeks after the beginning of the therapy; almost half (45.4%) had clinically improved at the end of the treatment; and at the posttreatment follow-up 39.7% fulfilled the criteria for clinical improvement.

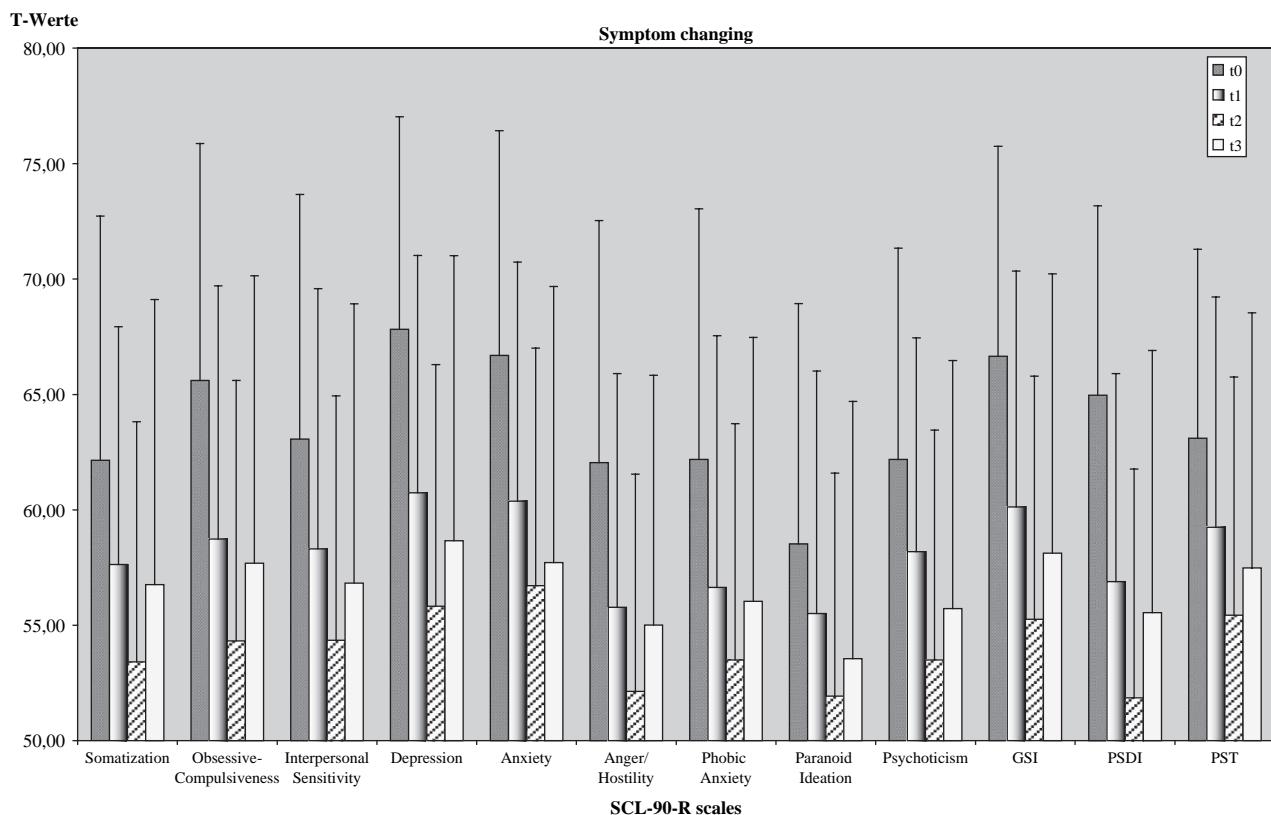


Figure 1. Change of psychological distress (Symptom Checklist-90-R (SCL-90-R)) over chronological sequence of therapy ( $N=156$ ,  $t_0$  = beginning of therapy,  $t_1$  = 4 weeks later,  $t_2$  = end of treatment,  $t_3$  = 1-year catamnesis).

Levels of symptom strain remained unchanged for 37.2% of participants 4 weeks after onset of therapy, for 22.4% at the date of dismissal, and for 26.9% at the posttreatment follow-up. From the beginning of therapy up to the intermediate measurement, six patients (3.9%) evinced a clinically significant deterioration, as did three patients (1.9%) at the date of dismissal and two patients (1.3%) by the posttreatment follow-up.

### Interpersonal Problems

Figure 2 demonstrates that, on average (stanine medium  $M=5 \pm 1 SD$ ), patients' values were located in the normative range on all IIP-C scales at the beginning of the therapy (stanine values based on ipsative scale values). In the course of treatment and at the posttreatment follow-up, the patients remained within the normative range. Comparison at

Table I. Symptom Checklist-90-R: Analyses of Variance Over Chronological Sequence of Therapy From the Beginning of Therapy ( $t_0$ ) Up to the Catamnesis ( $t_3$ )

Scales	F (VA%)	Partial $\eta^2$	$d_{t0-t1}$	$d_{t1-t2}$	$d_{t0-t2}$	$d_{t0-t3}$
Depression	70.74**	31.30	0.73	0.47	1.22	0.84
Obsessive-Compulsiveness	57.26**	27.00	0.65	0.40	1.05	0.69
Anxiety	54.59**	26.00	0.63	0.35	1.00	0.82
Anger/Hostility	47.79**	23.60	0.61	0.37	1.00	0.66
Psychoticism	46.39**	23.00	0.43	0.49	0.91	0.65
Phobic Anxiety	38.88**	20.10	0.51	0.30	0.82	0.55
Interpersonal Sensitivity	37.98**	19.70	0.43	0.36	0.82	0.55
Somatization	37.85**	19.60	0.43	0.41	0.83	0.47
Paranoid Ideation	25.12**	13.90	0.29	0.35	0.66	0.46
GSI	69.10**	30.80	0.67	0.47	1.16	0.80
PSDI	87.16**	36.00	0.89	0.50	1.41	0.97
PST	40.08**	20.50	0.45	0.40	0.84	0.57

Note. VA = inference for variance from partial  $\eta^2$ .

\* $p < .01$ . \*\* $p \leq .001$ .

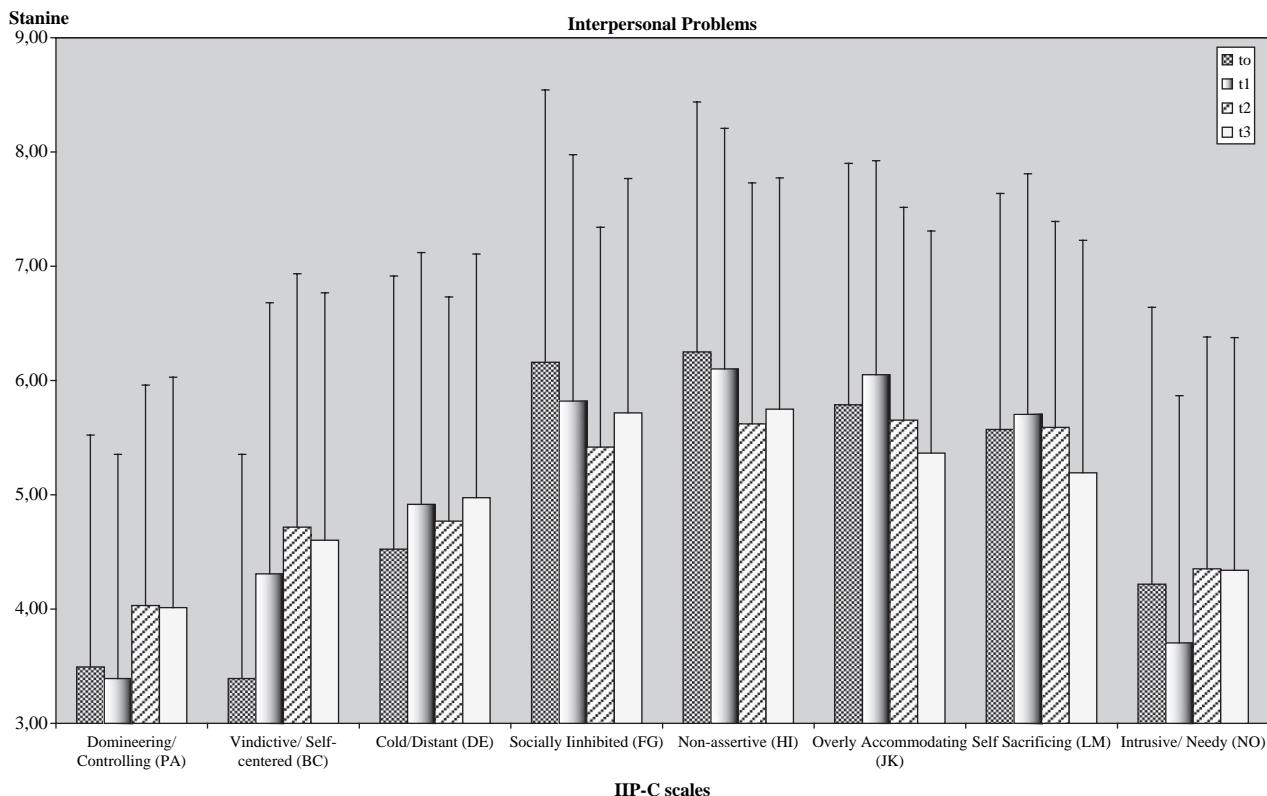


Figure 2. Interpersonal problems (Inventory of Interpersonal Problems Composite [IIP-C]) over chronological sequence of therapy ( $N = 156$ ,  $t_0$  = beginning of therapy,  $t_1 = 4$  weeks later,  $t_2$  = end of treatment,  $t_3$  = 1-year catamnesis).

mean value levels revealed only slight changes in the field of interpersonal problems.

Table II depicts statistically significant changes recorded on the IIP-C with the aid of ANOVA over the course of treatment (MANCOVA with repetition of measurement over all eight scales),  $F(3, 465) = 26.2$ ,  $p < .0001$ , 14.5% calculated VR. The extent of VR was highest in the course of treatment on the scales Vindictive/Self-Centered (BC) at 11.90% VR,  $F(3, 465) = 20.92$ ,  $p < .0001$ , and Domineering/Controlling (PA) with an explained variance of 6.90%,  $F(3, 465) = 11.42$ ,  $p < .0001$ ; VR was lowest

on the scale for Self Sacrificing (LM) at 2.40%,  $F(3, 465) = 3.8$ ,  $p < .01$ .

The Bonferroni-corrected paired individual comparisons undertaken between the beginning of the therapy ( $t_0$ ) and 4 weeks later ( $t_1$ ) yielded significant differences ( $p < .0001 - .009$ ) only on the scales BC and Intrusive/Needy (NO). Significant differences were found on the PA, BC, Socially Inhibited (FG), as well as Nonassertive (HI;  $p < .001 - .0001$ ) scales between the point of admission ( $t_0$ ) and the point of dismissal ( $t_2$ ). At posttreatment follow-up, significant differences appeared on the PA, BC, and

Table II. Inventory of Interpersonal Problems Composite: Analyses of Variance Over Chronological Sequence of Therapy From the Beginning of Therapy ( $t_0$ ) Up to the Catamnesis ( $t_3$ )

Scale	$F$ (VA%)	Partial $\eta^2$	$d_{t0-t1}$	$d_{t0-t2}$	$d_{t0-t3}$	$d_{t2-t3}$
Vindictive/Self-Centered (BC)	20.92**	11.90	-0.42	-0.63	-0.59	0.05
Domineering/Controlling (PA)	11.42**	6.90	0.05	-0.27	-0.26	0.01
Socially Inhibited (FG)	8.16**	5.00	0.15	0.34	0.20	-0.15
Intrusive/Needy (NO)	7.87**	4.80	0.22	-0.06	-0.05	0.01
Nonassertive (HI)	6.69**	4.10	0.07	0.29	0.24	-0.06
Overly Accommodating (JK)	6.24**	3.90	-0.13	0.07	0.21	0.15
Self-Sacrificing (LM)	3.80**	2.40	-0.06	-0.01	0.18	0.21
Cold/Distant (DE)	3.87**	2.40	-0.17	-0.11	-0.20	-0.1

Note. VA = inference for variance from partial  $\eta^2$ .

\* $p < .01$ . \*\* $p \leq .001$ .

Table III. Symptom Checklist-90-R: Mean Global Severity Index (GSI) Over Chronological Sequence of Therapy From the Beginning of Therapy ( $t_0$ ) Up to the Catamnesis ( $t_3$ )

Scale	F	Partial $\eta^2$ (VA%)	GSI <sub>t0</sub> (SD)	GSI <sub>t1</sub> (SD)	GSI <sub>t2</sub> (SD)	GSI <sub>t3</sub> (SD)
Inventory of Interpersonal Problems Composite changes ( $t_0 - t_2$ )						
Clinically and statistically significant improvements ( $n = 70$ )	4.91*	8.80	67.86 (9.31)	58.74 (10.49)	51.81 (10.05)	55.57 (11.97)
Normal range ( $n = 57$ )			63.60 (8.48)	58.05 (9.04)	54.46 (8.33)	56.70 (11.09)
Strain remained unchanged ( $n = 19$ )			69 (8.22)	68.58 (8.51)	66.84 (9.39)	68.37 (9.20)
Significant deterioration ( $n = 10$ )			67.90 (8.52)	65.70 (8.65)	62.00 (9.95)	64.60 (12.55)

Note. VA = inference for variance from partial  $\eta^2$ .

\* $p < .01$ . \*\* $p \leq .001$ .

Nonassertive (HI;  $p < .0001 - .027$ ) scales. Between dismissal ( $t_2$ ) and the 1-year follow-up ( $t_3$ ), no statistically significant changes could be ascertained on IIP-C scales.

### Effect Size

The IIP-C stanine values of the sample showed only mild effects when comparing the measurements at  $t_0$ ,  $t_1$ ,  $t_2$ , and  $t_3$ . Only the scale Overly Competitive showed moderate effects over the course of treatment ( $d_{t0-t1} = -0.53$ ,  $d_{t0-t2} = -0.72$  and  $d_{t0-t3} = -0.61$ ).

### Clinical and Statistical Significance

Following the model developed by Jacobson and Truax (1991), 70 (44.9%) patients showed clinically or statistically significant improvements at the end of treatment in the area of interpersonal problems. Fifty-seven (36.5%) were in the normal range of the IIP-C at admission and evinced neither a statistically significant improvement nor deterioration at dismissal. In addition, 10 (6.4%) patients fulfilled the criteria for a clinically significant deterioration and 19 (12.2%) for a statistical deterioration.

Consequently, we classified four different groups of patients. The first group includes patients who showed clinically and statistically significant improvements ( $n = 70$ ) in the area of interpersonal problems over the course of therapy. Patients scoring in the normal range of the interpersonal problems scales ( $n = 57$ ) over the course of therapy formed the second group. The third and fourth groups consisted of patients who either showed significant deterioration in interpersonal relations ( $n = 10$ ) or who, from entry to point of dismissal, still showed noticeable and unchanged departures from the norm ( $n = 19$ ).

The MANCOVA of the GSI, used over the course of therapy up to the posttreatment follow-up and taking into account the four groups in relation to the IIP-C outcome, yielded a statistically significant modification over the temporal course of treatment

as well during the follow-up period ( $GSI_{t0} - GSI_{t3}$ ) and comprised various modifications within the groups,  $F(3, 465) = 6.37$ ,  $p < .006$ .

The two patient groups whose interpersonal problems either worsened,  $F(3, 27) = 1.99$ ,  $p < .14$ , or remained unchanged but within the field of noticeable symptoms,  $F(3, 54) = 1.73$ ,  $p < .17$ , showed no statistically significant improvement in symptom strain during the course of therapy up to the posttreatment follow-up.

### Discussion

Inpatient and outpatient settings are marked by considerable differences. These concern, above all, the period of time in which the therapeutic dose is applied and the fact that, in inpatient therapy, not one therapeutic method but rather a mixture of different methods in combination is used. However, if one strictly applies the dose-effect model of Howard et al., then both arguments can be said to lose force, because this model presumes, first, that the dose—not the temporal factor—is decisive for the therapeutic effect and, second, that general and independent success factors are of greater importance to various schools than those factors that are specific to the respective therapeutic orientation. In the spirit of the model, then, inpatient psychotherapy can serve as a virtual paradigm for research into the effectiveness of various therapeutic doses (including those  $> 25$  hr) under largely standardized conditions.

If our results and those of other studies cited previously on the effectiveness of inpatient psychotherapy are integrated into the phase model developed by Howard et al. (1993), then the first week of treatment serves the purpose of alleviation and improvement of the patient's sense of well-being (remoralization). This effect emerges already after only a few days (eight therapeutic sessions). The improvement of symptoms (remediation) can be expected within the first 3 weeks in the framework of the therapeutic dose we applied (9–26 sessions).

The findings of our study confirmed the model insofar as the patients experienced a marked improvement in symptoms after a therapeutic dose of 40 hr (4 weeks). Their improvement within this period of time was significant. However, to restore earlier levels of functioning or to promote the learning of new roles (rehabilitation), a dose of 40 hr was obviously not enough. This could be clearly demonstrated in relation to interpersonal problems with the help of the IIP-C (Horowitz et al., 2000), because changes within the first 40 hr of therapy (4 weeks) were calculated only in the fields BC and NO and were thus classified as slight. Only at the point of dismissal from the clinic (after 8.6 weeks; i.e., after an average therapeutic dose of more than 80 hr) were significant changes observed. On the BC and PA scales, a constant increase in low stanine values could be ascertained at the end of therapy. Our interpretation is that patients learned in the course of therapy to assert themselves more successfully or to shrink from confrontations much less than they did at the beginning of treatment. Furthermore, until the time of dismissal there was a continuous decrease on the FG, Nonassertive (HI), and Overly Accommodating (JK) scales. These results indicate that the intervention of inpatient psychotherapy, carried out on average between Session 40 and the end of therapy, can be considered a rehabilitative phase in the sense of the phase model, whereas remoralization and remediation can apparently be achieved already after a dose of 40 therapeutic hours.

In addition to the aspect of dose–effect relationships, we considered the construct of clinical significance (Schmitz & Davies-Osterkamp, 1997). Patients who fail to show statistical or clinical improvement in the area of interpersonal problems from the onset of therapy to the end of treatment likewise evince no clinically relevant modification of their symptomatic complaints. Therefore, a deterioration in interpersonal problems during the course of therapy can be considered an indicator for the nonimprovement of symptom-related problems. Nevertheless, the improvement rate of participants over the course of therapy was not only statistically significant but also clinically relevant in 45% of cases up to the end of treatment. Our pre–post comparison of the sample showed a profile of clinically significant results similar to those in the study by Franz et al. (2000), who reported that 37% of their patients met the criteria for clinical significance. However, we were unable to achieve the 60.9% found in our own preliminary study of a smaller sample. From the point of admission up to the end of treatment, 23% of patients showed no statistically significant change in the area of symptoms. One part of this group (5%) comprised the

“normal testers” (clinically nonconspicuous as a result of certain test results). Interpreting the findings for these patients is difficult because it is possible that we are dealing with people who have perceptual difficulties in the face of their own psychic strain. For such patients, a successful treatment might involve learning to recognize their own symptomatic complaints better. Schmitz and Davies-Osterkamp (1997) consider the possibility that the extent of these patients’ symptomatic complaints might not be readable on the SCL-90-R (Franke, 2002).

When comparing the effect sizes found in our investigation with those of similar posttreatment follow-up studies (Beutel et al., 2005; Franke, Hoffmann, & Frommer, 2005; Sack, Lempa, Lampecht, & Schmidt-Ott, 2003), it becomes obvious that patients in these studies had a strong symptom reduction at the end of psychotherapy, reflected by high effect sizes ( $d_{GSI} = 1.11 - 1.22$ ). One study mentioned previously (Franz et al., 2000) also showed high effect sizes ( $d_{GSI} = 0.84$ ) at the end of treatment, albeit within a narrower scope. Reasons for this can be found, on the one hand, in the higher number of severely chronic patients; on the other hand, this may also be due to an interpretive effect of data selection among the just-mentioned studies of posttreatment follow-ups (Franke et al., 2005; Junge & Ahrens, 1996; Sack et al., 2003). The conclusion that those patients who did not participate in the follow-ups (dropout rates of 38–62%) did not benefit to such a great extent from inpatient psychotherapy seems unavoidable. We will follow up on this interesting problem in a forthcoming study. At the 1-year mark after treatment, a modest decline in symptomatic improvement could be observed, manifested in moderate to high effect sizes ( $d_{GSI} = 0.80$ ). Nevertheless, the patients indicated a significantly lesser symptom strain 1 year after treatment than at the beginning of treatment. This effect is known from other studies (Beutel et al., 2005; Junge & Ahrens, 1996) and cannot be overlooked, because it belongs to the patients’ developmental process. In terms of group statistics, it can be said that the participants at the time of the posttreatment follow-up indeed did not show the same level of symptom relief as at the end of therapy but had reached a level similar to that of the intermediate measurement 4 weeks after the beginning of therapy.

However, the inclusion of such a heterogeneous group of patients limits the value of these important conclusions. Thus, for example, the patient group shows considerable variations not only in the areas of interpersonal problems as well as symptom strain at the beginning of therapy but also in terms of therapeutic indications (e.g., ICD-10 diagnoses).

Aside from the heterogeneity of this patient group, the high number of dropouts distinguishing themselves in statically significant ways in terms of both educational level and age from the follow-up group could have resulted in a distortion of the therapeutic effects. Furthermore, there are other possible factors during the follow-up period, especially outpatient psychotherapeutic treatment or critical life events not considered in this study, that can complicate the results already identified. Such results cannot be attributed purely and exclusively to the treatment program described previously. The lack of a randomized control group must be counted as a limitation in interpreting the results. Consequently, concrete conclusions regarding an adequate length of treatment for psychotherapeutic inpatients are not yet possible and require further studies.

In spite of these limitations, some preliminary recommendations for inpatient psychodynamic treatment can be made. Not only this study but also other long-term investigations (Mohr, 1995) make clear that stasis or deterioration in the field of interpersonal problems do not amount to a long-term effect of psychotherapy. To be sure, the manifestation of real changes in interpersonal problems requires a longer therapeutic treatment than the amelioration of merely symptomatic distress.

### Note

$$^1 d = \frac{M_1 - M_2}{\sqrt{(s_1^2 + s_2^2)/2}}$$

### References

- Bachrach, H. M., Galatzer-Levy, R., Skolnikoff, A., & Waldron, S. (1991). On the efficacy of psychoanalysis. *Journal of the American Psychoanalytic Association*, 39, 871–916.
- Barkham, M., Rees, A., Stiles, W. B., Hardy, G. E., & Shapiro, D. A. (2002). Dose-effect relations for psychotherapy of mild depression: A quasi-experimental comparison of effects of 2, 8, and 16 sessions. *Psychotherapy Research*, 12, 463–474.
- Beutel, M. F., Höflich, A., Kurth, R., Brosig, B., Gieler, U., Milch, W., & Reimer, C. H. (2005). Short-term and long-term inpatient psychotherapy: Indication, results, predictors. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 51(2), 145–162.
- Bräutigam, W., Senf, W., & Kordy, H. (1990). Wirkfaktoren stationärer psychoanalytischer Therapien aus der Sicht des Heidelberger Katamnesenprojektes [Effects of analytic psychotherapy from the point of view of the Heidelberg Catamnestic Project] In H. Lang (Ed.), *Wirkfaktoren der Psychotherapie* [Effects of psychotherapy] (pp. 189–208). Berlin: Springer.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Davies-Osterkamp, S., Strauss, B. M., & Schmitz, N. (1996). Interpersonal problems as predictors of symptom-related treatment outcome in long-term psychotherapy. *Psychotherapy Research*, 6, 164–176.
- Derogatis, L. R. (1977). *SCL-90-R, administration, scoring and procedures manual-I for the R(revised) version*. Baltimore, MD: Johns Hopkins University School of Medicine.
- Franke, G. H. (2002). *Die Symptom-Checkliste von Derogatis—Deutsche version—Manual, ed 2, revised* [Symptom Checklist of Derogatis—German version—Manual, 2nd ed., revised]. Göttingen, Germany: Beltz.
- Franke, G. H., Hoffmann, T., & Frommer, J. (2005). Does improvement of symptoms four weeks after the beginning of psychodynamic inpatient psychotherapy correspond to long term outcome? *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 51, 360–372.
- Franz, M., Janssen, P., Lensche, H., Schmidtke, V., Tetzlaff, M., Martin, K., et al. (2000). On the effects of psychoanalytic oriented psychotherapy—An inpatient multicenter study. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 46, 242–258.
- Grissom, G. R., Lyons, J. S., & Lutz, W. (2002). Standing on the shoulders of a giant: Development of an outcome management system based on the dose model and phase model of psychotherapy. *Psychotherapy Research*, 12, 397–412.
- Hessel, A., Schumacher, J., Geyer, M., & Brähler, E. (2001). Symptom-Checkliste SCL-90-R. Testtheoretische Überprüfung und Normierung an einer bevölkerungsrepräsentativen Stichprobe [Validation and standardization based on a representative sample of the German population]. *Diagnostica*, 47, 27–39.
- Heuft, G., Eich, W., Henningsen, P., Janssen, P. L., Merkle, W., Fichter, M., et al. (2002). Psychosomatic and psychotherapeutic medicine goes to DRG—Prozedurenkatalog OPS 301 2.1 als erster Schritt. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 48, 90–103.
- Horowitz, L. M., Strauß, B., & Kordy, H. (2000). *Inventar zur Erfassung interpersonaler Probleme* (Deutsche version, ed 2, revised) [Inventory of interpersonal problems-German version, ed 2, revised]. Weinheim, Germany: Beltz.
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1989). The dose-effect relationship in psychotherapy. *American Psychologist*, 41, 159–164.
- Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology*, 61, 678–685.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Junge, A., & Ahrens, S. (1996). Inpatient psychosomatic treatment—Patient markers and treatment success. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 46, 430–437.
- Klose, C., Matteucci-Gothe, R., & Linden, M. (2006). Pre- and post-treatment in inpatient psychosomatic rehabilitation. *Rehabilitation*, 45, 359–368.
- Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 159–172.
- Lambert, M. J., & Ogles, B. M. (2004). The efficacy and the effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., rev., pp. 139–193). New York: Wiley.
- Lang, T., & Hoyer, J. (2003). Klinische Untersuchungsverfahren. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 32(4), 332–334.
- Leichsenring, F. (2005). Are psychodynamic and psychoanalytic therapies effective? A review of empirical data. *The International Journal of Psycho-Analysis*, 86, 841–868.

- Lueger, R. J. (1995). Ein Phasenmodell der Veränderung in der Psychotherapie. *Psychotherapeut*, 40, 267–278.
- Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice*, 2, 1–27.
- Rudolf, G., Jakobsen, T., Micka, R., & Schumann, E. (2004). Results of psychodynamic inpatient psychotherapy in relation to diagnosis. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 50, 37–52.
- Sack, M., Lempa, W., Lamprecht, F., & Schmidt-Ott, G. (2003). Individualized therapy goals and treatment efficacy—A follow-up study one year after inpatient psychotherapy. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 49, 63–73.
- Schmitz, N., & Davies-Osterkamp, S. (1997). Klinische und statistische Signifikanz—Diskutiert am Beispiel der Symptom Check Liste (SCL-90-R) [Clinical and statistical significance using the symptom check list (SCL-90-R)]. *Diagnostica*, 43(1), 80–96.
- Stiles, W. B., & Shapiro, D. A. (1989). Abuse of the drug metaphor in psychotherapy process-outcome research. *Clinical Psychology Review*, 9, 521–543.
- Wuchner, M., Eckert, J., & Biermann-Ratjen, E. M. (1993). Vergleich von Diagnosegruppen und Klienten verschiedener Kliniken. *Gruppenpsychotherapie Gruppendynamik*, 29, 198–214.

**Copyright of Psychotherapy Research is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.**