Psychoanalytic Versus Psychodynamic Therapy for Depression: A Three-Year Follow-Up Study

Dorothea Huber, Gerhard Henrich, John Clarkin, and Guenther Klug

The aim of this study was to investigate the effectiveness of long-term psychoanalytic and psychodynamic psychotherapies. In a prospective, randomized outcome study, psychoanalytic (mean duration: 39 months, mean dose: 234 sessions) and psychodynamic (mean duration: 34 months, mean dose: 88 sessions) therapy were compared at post-treatment and at one-, two-, and three-year follow-up in the treatment of patients with a primary diagnosis of unipolar depression. All treatments were carried out by experienced psychotherapists. Primary outcome measures were the Beck Depression Inventory and the Scales of Psychological Capacities, and secondary outcome measures were the Global Severity Index of the Symptom Checklist 90-R, the Inventory of Interpersonal Problems, the Social Support Questionnaire, and the INTREX Introject Questionnaire. Interviewers at pre- and post-treatment and at one-year follow-up were blinded; at two- and three-year follow-up, all self-report instruments were mailed to the patients. Analyses of covariance, effect sizes, and clinical significances were calculated to contrast the groups. We found significant outcome differences between treatments in terms of depressive and global psychiatric symptoms, personality functioning, and social relations at three-year follow-up, with psychoanalytic therapy being more effective. No outcome differences were found in terms of interpersonal problems. We concluded that psychoanalytic therapy associated with its higher treatment dose shows longer-lasting effects.

The effectiveness of short-term psychodynamic psychotherapy for depressive disorders is well established, as a recent meta-analysis has confirmed (Driessen et al., 2010). But with accumulating scientific evidence of the recurrent nature of unipolar depression in the last decade, psychotherapy outcome research has evidenced a shift away from assessing simply whether treatment leads to recovery and remission of acute symptoms toward studies based on the more fine-grained mechanisms of depression

Dorothea Huber, Ph.D., M.D. Gerhard Henrich, Ph.D., and *Guenther Klug, M.D.,* are all affiliated with the Klinik und Poliklinik für Psychosomatische Medizin und Psychotherapie, Klinikum rechts der Isar, Technische Universität München, in Munich, Germany. Dr. Huber is also affiliated with the International Psychoanalytic University Berlin.

John Clarkin, Ph.D., is with Weill Cornell Medical College, in New York City.

The study was supported by grants from the Deutsche Forschungsgemeinschaft (DFG; 444 USA 111/ 3 / 98); the Research Advisory Board of the International Psychoanalytic Association; and the Dr. Zita und T. V. Steger-Stiftung.

Address correspondence to Professor Dorothea Huber, Ph.D., M.D., Klinik für Psychosomatische Medizin und Psychotherapie, Klinikum München-Harlaching Sanatoriumsplatz 2, 81545 München, Germany. E-mail: D.Huber@lrz.tu-muenchen.de

(Luyten & Blatt, 2011), examining whether treatment may prevent future symptom recurrence or a chronic illness course. Thus, the effectiveness of a treatment cannot be evaluated adequately by measuring its influence on an index episode; the litmus test rather is the prevention of recurrence and relapse. When scrutinizing the results of short-term psychotherapy, at least for Major Depressive Disorder (MDD), residual symptoms and recurrence seem to be quite frequent, and chronic cases are not treated sufficiently. Studies show that residual symptoms, recurrence, and relapse are common after short-term treatment for depressive disorders (Fava, Ruini, & Belaise, 2007; Koppers, Peen, Niekerken, Van, & Dekker, 2011; Taylor, Walters, Vittengl, Krebaum, & Jarrett, 2010) and that chronic cases are insufficiently treated (Cuijpers et al., 2010; Dunner, 2001). Some argue that the best solution to this problem is to provide maintenance and continuation therapy. Others argue that short-term treatments are insufficient (Shea et al., 1992) and advocate for the provision of long-term treatments as a more efficient and potentially cost-effective way to help patients. One proposed solution is to provide maintenance and continuation therapy, with on-going clinical monitoring with flexible frequencies and patterns (e.g., Vittengl, Clark, & Jarrett, 2009). Another proposed solution is to address the insufficient effectiveness of short-term treatment by providing long-term treatments as a potentially more efficient and cost-effective way to help patients to cope better with potential future stressors likely to trigger depression. Although the effectiveness of long-term psychotherapy for general mental disorders has been demonstrated in several studies (Grande et al. 2006; Jakobsen et al., 2007; Knekt et al., 2008; Knekt et al., 2011; Leuzinger-Bohleber, Stuhr, Rueger, & Beutel, 2003; Wallerstein, 1986), one review (De Maat, de Jonghe, Schoevers, & Dekker, 2009), and two meta-analyses (Leichsenring & Rabung, 2008; Leichsenring & Rabung, 2011), well-designed studies of long-term treatment especially for depressed patients are still lacking.

Beyond the question of whether therapy works, studies are beginning to examine the question of how and why it works (Kazdin, 2007). Little is known about mechanisms of therapeutic change in general and for MDD specifically, but there is growing empirical evidence that impaired personality functioning leads to social maladjustment, dysfunctional attitudes, and faulty attributions that in turn predispose patients to relapse and recurrence (Fava et al., 2007). Research on mechanisms of change suggests that psychotherapies should target personality functioning in order to efficiently go beyond symptom relief, but empirical findings are only tentative and knowledge is sparse as to which treatment models and which parameters (such as treatment duration and session frequency) help patients achieve and maintain change in personality functioning.

Since research evidence on the recurrent nature of unipolar depression has accumulated in the latest decade (Judd, 1997), psychotherapy outcome research has evidenced a paradigm shift away from assessing mere recovery from acute symptoms toward examining the prevention of relapse, recurrence, and chronicity of the disorder. Thus, the emerging consensus is that the effectiveness of a treatment cannot be evaluated adequately by assessing its effect on an index episode; the litmus test rather is the stability of its effect over a sufficiently long follow-up period. To estimate an adequately long follow-up period, researchers must consider the natural course of the disorder. In the last decades, epidemiologic research has provided valid data for an estimation of the natural course of depression. According to Eaton's (Eaton, 2008) summary of the Baltimore Epidemiologic Catchment Area (ECA) Follow-up, a population-based sample, as many as 53% of patients with a first lifetime episode of depressive disorder either do not recover at all or have at least one recurrence. The National Institute of Mental Health Collaborative Study of the Psychobiology

of Depression (Keller et al., 1992; Keller & Shapiro, 1981) reported from a special psychiatry setting a probability of recurrence of 67% after 10 years and 85% after 15 years. A consistent finding among different studies is that the risk of recurrence increases after each subsequent episode. Therefore, Roth and Fonagy (2005) recommend a two-year follow-up, while Frank and colleagues (1990) recommend a three-year follow-up to disentangle treatment effects from the natural course of the disorder. To address the question of long-term treatment effects for depression, we investigated the effectiveness of two long-term treatments (psychoanalytic and psychodynamic therapy; for a detailed operationalization, see below, p. 138) with a three-year follow-up. We investigated "real-world" treatments provided in the German health-care system, using a design balancing the concerns of both internal and external validity. We hypothesized that psychoanalytic therapy would be significantly superior to psychodynamic therapy in terms of both (a) symptom improvement and (b) personality change.

METHODS

Study Design

The Munich Psychotherapy Study (MPS) is a comparative randomized study of psychoanalytic and psychodynamic therapy. It is designed to maximize external validity by examining non-manualized and representative psychotherapies under the conditions of day-to-day practice conducted by experienced psychotherapists, while maximizing internal validity by recruiting a diagnostically homogeneous sample and randomizing participants to the two treatment groups. The study was conducted at the Department of Psychosomatic Medicine and Psychotherapy, Technische Universität München (TUM, Germany). The study protocol was approved by the Ethics Committee of the TUM.[1]

Participants

Participant flow is presented in Figure 1. Patients seeking treatment at the outpatient clinic completed the Beck Depression Inventory (BDI; Hautzinger, Bailer, Worall, & Keller, 1994). If BDI was at least 16, an intake interview was conducted by investigator I who used ICD-10 (WHO, 1993)/*DSM-IV* (APA, 1994) criteria to assess for depressive disorder and personality disorder. We screened 101 depressive patients with BDI > 16 in this manner.

Two psychiatrists consensually assessed the type of depressive disorder (and, if appropriate, any co-morbid axis I and II disorders) by means of the clinical interview that was recorded and subsequently reviewed by the two psychiatrists using the International Diagnostic Checklists for ICD-10 and *DSM-IV* (ICD-10/*DSM IV* checklists; Hiller, Zaudig, & Mombour, 1995). These checklists include diagnostic criteria to enable clinicians and researchers to reliably diagnose mental disorders. Interrater reliability (kappa) was .70 for depressive episode and .72 for recurrent depressive disorder (Hiller et al., 1994). Thus, inclusion criteria required participants to have (a) a BDI total score of at least 16 and (b) a primary diagnosis of a major depressive disorder with a current moderate or severe episode (ICD-10 diagnoses F 32.1/2 or F 33.1/2 or *DSM-IV* diagnoses 296.22/23 or 296.32/33) or a double depression characterized by both dysthymic disorder and a current major depressive episode (Keller, Hirschfeld, & Hanks, 1997). Exclusion criteria included mild depressive episode, bipolar affective disorder, depression due to somatic illnesses or diseases of the brain, alcohol or substance dependence, psychotherapy during the past two years, and con-

1. We also included a CBT condition in our study, but did this after randomization of the two other groups was completed; the results of this quasi-experimental design will be presented later.

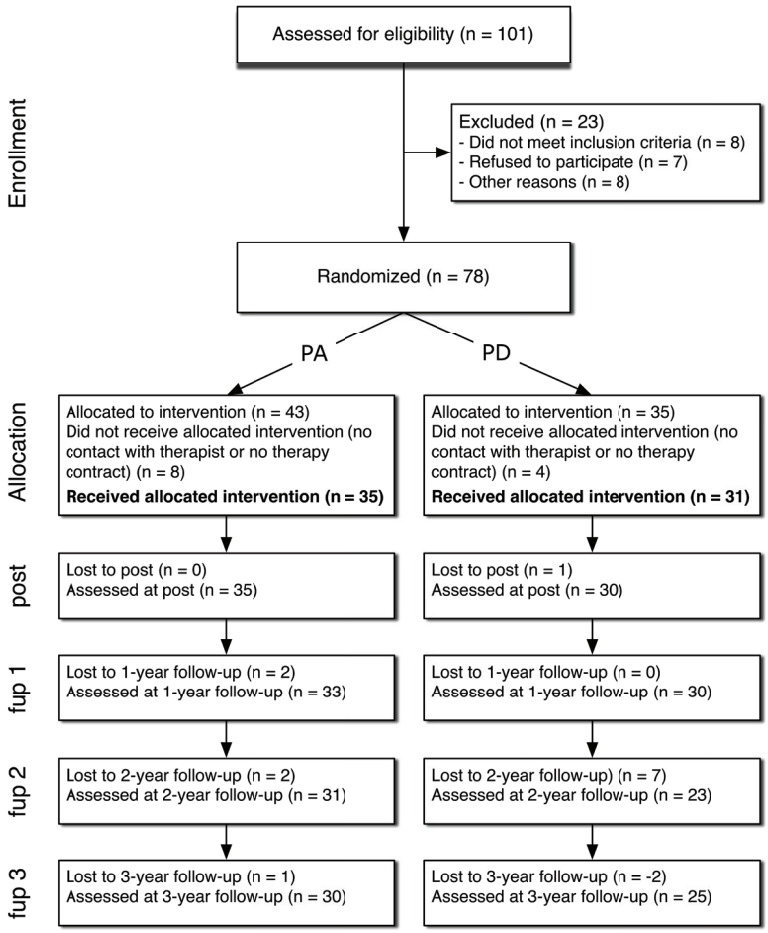

FIGURE 1. Participant flow.

current anti-depressive medication. In all, 23 patients were excluded because they failed to meet the study criteria or refused to participate in the study. Thus, 78 patients signed informed consent and were randomized to the two treatment groups. Twelve patients did not contact the therapist or did not enter into a therapy contract after the five trial sessions (a trial period of five sessions is standard practice in Germany). Thus, 15% of the sample did not start psychotherapy. As recommended by Lambert and Ogles (2004), only those patients signing therapy contracts were included in the study, and consequently 35 patients in the psychoanalytic and 31 patients in the psychodynamic group were followed up (even if they did not complete treatment). During the course of therapy, one patient in the psychodynamic group and no patients in the psychoanalytic group dropped out. During the entire follow-up period, five patients in each group were lost. Thus, at 3-year follow-up, the psychoanalytic group included 30 and the psychodynamic group included 25 patients. There was no statistically significant difference in dropout rate (Fisher's exact test, $p = .743$).

Assessment and Measures

A multimodal and multidimensional measurement strategy was applied. A key goal of the study was to go beyond symptom measurement. Therefore, in addition to symptom measures, we included measures of interpersonal problems and personality functioning as well.

Symptoms were assessed using patient self-report on the German versions of the BDI and the Global Severity Index (GSI) of the Symptom Checklist (SCL-90-R; Franke, 1995). We used the total scores of the BDI and the GSI, where higher values represent worse symptom severity. Self-reported remission from symptoms was operationalized as a total score of < 10 in the BDI, which is in accordance to the S3- and national healthcare guideline "Unipolar Depression" (DGPPN et al., 2009), the review of Williams and colleagues (Williams, Noël, Cordes, Ramirez, & Pignone, 2002), and the suggestion of Knekt and colleagues (Knekt et al., 2008; Knekt et al., 2011).

Personality functioning was assessed using expert ratings on the Scales of Psychological Capacities (SPC; Huber, Brandl, & Klug, 2004; Huber, Henrich, & Klug, 2005; Huber, Klug, & Wallerstein, 2006) and patient self-report on the positive introject scale of the INTREX Introject Questionnaire, short version 2.0 (Tress, 1993). The SPC tap a set of constructs derived from the core elements of psychic functioning according to psychoanalytic theory. They comprise the ego's dealing with needs and drives, its control and modulation of affects and impulses, its regulation of closeness to and distance from others, internalizing and externalizing propensities, stability of mood, and adequacy of norms and ideals. Above all, the SPC assess problematic personality traits assumed to be the characterological basis for clinical disorders (Boyce, Parker, Barnett, Cooney, & Smith,, 1991; Krueger 1999), thus linking normal-range personality constructs to clinical disorders. These problematic personality traits are the less pervasive and more changeable dynamic elements of personality pathology; improvements during or after a treatment indicate a change in underlying dynamic elements of personality pathology. The SPC are an expert-rated measure to gauge changes in personality functioning by repeated measurement. The assessment is based on a one-hour clinical intake interview together with a one- to one and a half–hour semi-structured SPC interview. An extensive manual is available for the rating procedure (Huber, Klug, & Wallerstein, 2006). The material is rated on a seven-point scale from 0 for normal or fully adaptive functioning to 3 for seriously disturbed functioning, with half-points in between. The application of the instrument requires rater training to reach high agreement with a calibrated set of judgments by expert judges. Research groups from different sites (DeWitt, Milbrath, & Wallerstein, 1999; Huber et al., 2004; Huber et al., 2005) found independently satisfying psychometric qualities of the measure. In this study, each interview was rated by two raters. Mean interrater-reliability was 0.82. In a previous study, we found no significant correlation between the SPC and depressive complaints at pre-treatment (Huber et al., 2004). For statistical calculations, the mean of all scales (total score) was used. Detailed information of depression-specific SPC-sub-dimensions is given in Klug and Huber (2009).

The positive introject scale of the INTREX Introject Questionnaire (Tress, 1993), an 8-item self-rating instrument derived from the Structural Analysis of Social Behavior (SASB; Benjamin, 1974; Benjamin, 1983), was applied to grasp another self-reported dimension of personality functioning. The introject, one of the three surfaces of the SASB model, is a "hypothesized personality structure, which comprises a relatively stable conscious and unconscious repertoire of ways of treating the self" (Henry, Schacht, & Strupp, 1990, p. 769), including self-appraisals (e.g. self-accepting, self-nurturing, self-helping), verbal, and motor activities di-

rected at the self and self-images. We used the total score, where higher values represent more positive acts toward the self.

Social relations were assessed with the German version of the Inventory of Interpersonal Problems, 64-item short version (IIP; Horowitz, Strauß, & Kordy, 2000; Huber, Henrich, & Klug, 2007), using the total score with high values indicating more interpersonal problems, and the Social Support Questionnaire (F-SOZU short version; Sommer & Fydrich, 1991). The Social Support Questionnaire is a 22-item self-report questionnaire for the assessment of perceived social support. It comprises three dimensions: emotional support, practical support, and social integration. We used the total score where higher values represent more perceived social support. Psychometric qualities (reliability, content, and construct validity) are satisfactory (Dunkel, Antretter, Fröhlich-Walser, & Haring., 2005; Fydrich, Geyer, Hessel, Sommer, & Brähler, 1999).

The procedure was as follows: Patients presented to the clinic, completed the BDI, and were scheduled for an intake interview if BDI was 16 or higher. If the patient was deemed depressed according to the inclusion criteria and he/she signed the consent form, the SPC interview was conducted during a second session where the patient filled in the SCL-90-R, the IIP, F-SOZU, and INTREX introject positive. After that, the patient was randomly assigned. Patients had to wait between two and four weeks from randomization till beginning of treatment.

Measurement points were pre-treatment, post-treatment, and one-year, two-year, and three-year follow-up. Patients completed the SCL-90-R, BDI, IIP, F-SOZU, and INTREX introject positive at pre- and post-treatment and at one-, two-, and three-year follow-up. Investigator I, using a clinical interview, assessed patients with the ICD-10/*DSM-IV* checklists and the SPC interview at pre-treatment. Investigator II, who was blind to pre-treatment data, assessed the patients with the ICD-10/*DSM-IV* checklists, the clinical interview and the SPC interview at post-treatment and at one-year follow-up. Thus, investigators at all measurement points were blind to treatment modality, but not to measurement point.

The investigators were postgraduate physicians or psychologists at advanced stages of their therapeutic training, who had several years of clinical experience. Regular training and recalibration sessions were carried out. They had no access to information about patients, and patients were instructed to give no clues about the treatment. At two-year and at three-year follow-up, all self-report instruments were mailed to the patients, so we have no observer-rated data (e.g., no SPC data) for these measurement points.

As they are of specific importance for depressive disorders, the BDI and the SPC were specified a priori as primary outcome measures. Secondary outcome measures were SCL-90-R-GSI, IIP, F-SOZU, and INTREX introject positive.

Treatments and Therapists

The 14 study therapists were thoroughly trained according to the German Psychotherapy Guidelines (Rueger, Dahm, & Kallinke, 2005) in their analytic training institutes and approved by the German Association for Psychoanalysis, Psychotherapy, Psychosomatics and Depth Psychology (DGPT), the umbrella organization for psychoanalytic therapy schools in Germany. All therapists were highly experienced; mean duration of psychotherapeutic practice was 18 years (range: 8–29 years). Mean age was 49 years (range: 39–56 years). No school of psychoanalysis was mandatory. The main theoretical orientation—evaluated with the Therapeutic Attitude Questionnaire (ThAt; Klug, Henrich, Kaechele, Sandell, & Huber, 2008)—of nine therapists was a mixture of classical (Freudian) psychoanalysis and object-relational psychoanalysis, and five therapists were object relational psychoanalysts. Treatments were not formally manualized, and we decided to refrain from any kind of

supervision or competence checks since all therapists were already highly trained and experienced. The therapeutic modality, not the therapist, was randomly assigned to the patient in order not to interfere with the patient-therapist match. If therapist or patient decided at the beginning that they could not work with the other, the patient was sent back to the study center and referred to another therapist (of the same treatment condition). Patient and therapist decided together when to terminate treatment.

Psychoanalytic therapy is defined by Fonagy and Kaechele (2009, p. 1338) as a "predominantly verbal, interpretative, insight-oriented approach which aims to modify or re-structure maladaptive relationship representations that lie at the root of psychological disturbance." The usual duration of psychoanalytic therapy according to the German Psychotherapy Guidelines is 240 sessions; session frequency is 2–3 sessions/week with the patient lying on the couch. Psychodynamic therapy is based on the same principles of theory and technique, but it is more limited both in the depth of the therapeutic process and in its goals by focusing on symptom-sustaining here-and-now conflicts. The usual duration of psychodynamic therapy according to the German Psychotherapy Guidelines is 80 sessions. One weekly session is carried out in a face-to-face setting.

In the MPS, mean duration of psychoanalytic therapy was 39 months (range: 3–91) or 234 sessions (range: 17–370). Mean duration of psychodynamic therapy was 34 months (range: 3–108) or 88 sessions (range: 12–313). Thus, the two treatments were similar in duration but different in dose (number of sessions).

Treatment fidelity was assessed with both therapist- and expert-rated measures. Therapists completed the German version of the Periodical Rating Scale for Psychoanalytic Treatment (Beenen & Stoker, 1996) every 6 months, which comprises a manual with operationalizations of treatment parameters discussed with the therapists regularly. Based on the standard commentary of the German Psychotherapeutic Guidelines, we selected the following parameters to discriminate between psychoanalytic and psychodynamic treatment: Session frequency (0 = one hour/week, 1 = two hours/week, 2 = three hours/week), checked by date of session in therapists' insurance accounts; patient position (0 = sitting, 1 = couch); strength of transference on therapist (0 = weak, 1 = moderate, 2 = strong); and technique (0 = supportive, 1 = mixed supportive and insight-oriented, 2 = insight-oriented).

Each treatment was assessed by the therapist in this manner, and a total score was calculated. Higher scores indicate psychoanalytic format while lower scores indicate psychodynamic format, although in less thoroughgoing and less systematic ways psychodynamic therapy is partly insight-oriented, and effective psychoanalytic therapy applies supportive technique (Bush & Meehan, 2011; Gabbard & Westen, 2003; Kaechele, 2010). A two-tailed *t*-test revealed a significant difference in the expected direction between the psychoanalytic and the psychodynamic mean total rating scores ($t = 9.26$; $df = 55$; $p < .001$).

Independent raters assessed treatment fidelity using the Psychotherapy Process Q-set (PQS; Jones, 2000). The 100 items of the PQS capture key treatment processes, including patient behavior, therapist behavior, and patient-therapist interactions. We used the 20-item PQS prototype for psychoanalytic therapy described by Ablon and Jones (2005) to assess how much each treatment adhered to standard psychoanalytic practice. Fifty percent of all psychoanalytic treatments and fifty percent of all psychodynamic treatments were selected at random, and one audio-taped mid-phase session of each treatment was assessed by trained PQS raters blind to treatment modality. Mean scores of the 20 items of the prototype were calculated. A two-tailed *t*-test yielded a significant difference between the groups ($t = 3.2$, $df = 30$, $p < .01$).

Thus, according to the ratings of therapists and independent raters, the two treat-

TABLE 1. Sociodemographic and Clinical Characteristics of Both Treatment Groups Psychoanalytic (PA; n = 35) and Psychodynamic (PD; n = 31); for all diagnosis: first ICD-10, second *DSM-IV*.

	PA		PD		
Age	***M***	***SD***	***M***	***SD***	
years	31.2	5.6	34.9	8.0	$t(63) = 2.15$*
Gender	*n*	%	*n*	%	
female	24	69.0	18	58.1	$\chi^2(1) = 0.79$
Partnership	***n***	**%**	***n***	**%**	
in partnership	20	57.1	18	58.1	$\chi^2(1) = 0.01$
single	15	42.9	13	41.9	
Education	*n*	%	*n*	%	
final examination level	24	68.6	21	67.7	$\chi^2(1) = 0.01$
less than final examination level	11	31.4	10	32.3	
Employment	*n*	%	*n*	%	
full time	19	54.3	15	50.0	$\chi^2(1) = 1.76$
part time	2	5.7	4	13.3	
unemployed	2	5.7	2	6.7	
still in education	6	17.1	6	20.0	
housewife, senior	6	17.1	3	10.0	
Depression diagnosis	***n***	**%**	***n***	**%**	
Depressive episode F 32.1/2; Major Depression, single episode 296.22/3	15	42.9	16	51.6	$\chi^2(1) = 0.51$
Recurrent depressive episode F 33.1/2; Major Depression, recurrent 296.32/3	20	57.1	15	48.4	
double depression	21	60.0	15	48.4	$\chi^2(1) = 0.89$
Duration of depression	***M***	***SD***	***M***	***SD***	
months	72.5	81.8	58.8	82.5	$t(63) = 0.68$
Personality disorder diagnosis	*n*	%	*n*	%	
Yes	11	31.4	11	35.5	$\chi^2(1) = 0.12$
Schizoid personality disorder F 60.1; Schizoid personality disorder 301.20	3		3		
Emotionally unstable personality disorder, borderline type F 60.31; Borderline personality disorder 301.83	2		1		
Dependent personality disorder F 60.7; Dependent personality disorder 301.60	1		0		
Narcissistic personality disorder F 60.8; Narcissistic personality disorder 301.81	5		7		
Axis I comorbidity	*n*	%	*n*	%	
yes	13	37.1	11	35.5	$\chi^2(1) = 0.02$
Agoraphobia with panic disorder F 40.01; Agoraphobia (with a history of panic disorder) 300.21	2		1		
Social phobia F 40.1; Social phobia 300.23	0		1		
Specific isolated phobia F 40.2; Phobic disorders 300.29	1		1		
Panic disorder F 41.0; Panic disorder 300.01	2		1		
Somatization disorder F 45.0; Somatization disorder 300.81	1		1		
Undifferentiated somatoform disorder F 45.1; Undifferentiated somatoform disorder 300.82	1		0		
Somatoform autonomic dysfunction F 45.3	4		3		
Psychological and behavioral factors associated with disorders or diseases classified elsewhere F 54; Psychological factors affecting medical condition 316	2		3		
Prior treatment	*n*	%		*n*	%
yes	16	47.1	9	31.0	$\chi^2(1) = 1.68$
Inpatient treatment	6		5		
Outpatient treatment	10		4		

*$p < 0.05$; *df* corrected for heterogenous variances (Levene's test: $p < 0.05$)

ments were different in terms of important parameters and processes. The therapist ratings of technique and the independent PQS ratings were significantly correlated (Pearson $r = .58$, $n = 30$, $p < .001$), thus showing consistency between the two ratings of treatment differentiation.

Statistical Analysis

To concentrate on the long-term effectiveness, analyses of covariance (ANCOVAs) with pre-treatment values as covariates were conducted only at three-year follow-up. Self-reported remission from symptoms was operationalized as a total score of < 10 in the BDI; we tested significance with Fisher's exact test (two-tailed). Within- and between-subjects effect sizes (ES) were calculated by dividing pre-treatment/follow-up differences by the pooled pre-treatment standard deviation (within-subjects ES) or—according to Cohen's formula (Cohen, 1988)—dividing group differences at each measurement point by the appropriate pooled standard deviation (between-subjects ES). According to Cohen's benchmarks (which, in the strictest sense, are valid only for between-groups ESs), we differentiated between high (> .8), medium (> .5), and low (> .2) ES. In addition, clinical significance (CS) was calculated for the primary outcome variables (BDI, SPC) as the percentage of patients moving from the abnormal into the normal range (Jacobson & Truax, 1991). Data imputation in case of missing data was performed by the last observation carried forward procedure. A two-tailed alpha level of p = .05 was used to determine statistical significance; for the primary outcome variables, the alpha level was adjusted using a Bonferroni correction to $p = .025$.

Randomization was achieved by a computer-generated random number table. Sample size was based on a power-analysis, indicating that 63 subjects were needed for adequate power (= 0.8) to detect a large effect of f = 0.4 (equivalent to Cohen's d = 0.8), with an error probability alpha = .025 (for two primary outcome measures) for two groups and one covariate (pre-treatment values).

RESULTS

Patient characteristics are provided in Table 1. Except for age, we found no significant differences between the treatment groups in terms of socio-demographic and diagnostic data.

Table 2a presents means and standard deviations, Table 2b within-groups and between-groups effect sizes, and for the primary outcome variables indicators of clinical significance.

BDI scores showed large effects from pre- to post-treatment and to follow-up, indicating large improvements within both groups; within-group ESs were 2.3–2.6 for PA and 2.0–2.1 for PD, and at follow-up 3, as many as 91.4% of PA and 77.4% of PD patients were in the normal range. Between-group ES showed PA to be superior at each time-point, with an ES of 0.70 at follow-up 3, and a significant group difference at follow-up 3, $F(1, 63) = 8.43$, $p = .005$. Self-reported remission from depressive symptoms, operationalized as a total score of < 10 in the BDI, are significantly different (PA: 86% PD: 58%; Fisher's exact test, $p = .015$).

GSI scores showed large effects from pre- to post-treatment and to follow-up, indicating large improvements within both groups; within-group ESs were 1.5–1.8 for PA and 1.0–1.3 for PD. Between-group comparisons showed PA to be superior at each time point, with an ES of 0.60 at follow-up 3. We found significant group difference at follow-up 3, $F(1, 63) = 8.16$, $p = .006$).

TABLE 2A. Outcome Measure Scores at Pre-Treatment (Pre), Post-Treatment (Post), and 1-, 2- and 3-Year Follow-Up (fup 1 to fup 3) for Psychoanalytic (PA; $n = 35$) and Psychodynamic (PD; $n = 31$) Therapy

			pre		post		fup 1		fup 2		fup 3	
		N	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*
BDI	PA	35	25.49	7.31	6.57	5.96	7.06	6.06	6.74	6.21	4.60	4.06
	PD	31	25.10	8.68	8.26	9.89	9.23	8.38	9.39	10.74	9.52	9.29
SCL-90-GSI	PA	35	1.27	0.52	0.44	0.32	0.48	0.45	0.45	0.46	0.33	0.34
	PD	31	1.17	0.53	0.49	0.52	0.62	0.61	0.57	0.59	0.61	0.55
IIP	PA	35	14.37	3.42	9.59	4.63	9.32	4.66	9.36	4.36	8.34	4.68
	PD	31	14.82	2.53	10.46	4.11	11.64	4.00	11.18	3.95	9.99	4.49
F-SOZU	PA	33	3.75	0.65	4.15	0.73	4.25	0.53	4.25	0.57	4.31	0.61
	PD	30	3.74	0.78	4.00	0.78	3.96	0.72	3.92	0.75	3.87	0.89
SPC	PA	35	1.05	0.21	0.60	0.25	0.54	0.23				
	PD	31	1.07	0.26	0.77	0.31	0.71	0.31				
INTREX	PA	33	3.78	2.00	5.91	2.07	6.16	1.95	6.27	1.74	6.59	1.94
intr pos	PD	30	4.30	1.61	5.42	1.80	5.70	1.47	5.42	2.07	5.66	1.82

Note. BDI, Beck Depression Inventory; SCL-90-GSI, Symptom Checklist-90-R, Global Severity Index (GSI); IIP, Inventory of Interpersonal Problems; F-SOZU, Social Support Questionnaire, short version K-22; SPC, Scales of Psychological Capacities; INTREX Introject Questionnaire, short version 2.0, scale: introject positive; ES, effect size; CS, clinical significance

IIP total scores showed large effects from pre- to post-treatment and to follow-up, indicating large improvements within both groups; within-group ESs were 1.6–2.0 for PA and 1.0–1.6 for PD. Although the between-groups ESs were 0.4–0.5 (medium) for follow-ups 1–3, group comparison at follow-up 3 showed no significant differences between PA and PD, $F(1, 63) = 1.71$, $p = .196$.

F-SOZU scores indicated improvements from pre- to post-treatment and to follow-up within both groups; within-group ESs were 0.6–0.8 (medium to large) for PA, but only 0.2–0.4 (small to medium) for PD. Between-group comparisons showed PA to be superior at each time point, with a medium ES of 0.6 and a significant difference at follow-up 3 , $F(1, 60) = 8.72$, $p = .004$).

SPC scores showed large effects from pre- to post-treatment and to follow-up 1, indicating large improvements within both groups; within-group ESs were 1.9–2.2 for PA and 1.3–1.5 for PD, and at follow-up 1 (last available measurement), as many as 68.6% of PA and 38.7% of PD patients were in the normal range. Between-group comparisons showed PA to be superior at each time point, with a medium ES of 0.6 and a significant group-difference at follow-up 1, $F(1, 63) = 6.81$, $p = .011$.

INTREX scores indicated medium to large improvements from pre- to post-treatment and to follow-up within both groups; within-group ESs were 1.2–1.5 (large) for PA and 0.6–0.8 (medium) for PD. Between-group comparisons showed PA to be superior at each time point, with a medium ES of 0.5, and a significant difference at follow-up 3, $F(1, 60) = 6.30$, $p = .015$).

DISCUSSION

This study examined the effectiveness of long-term psychoanalytic and psychodynamic treatment for depression. We investigated the stability of treatment effects over

TABLE 2B. Effect Sizes and Clinical Significances at Post-Treatment (Post), and 1-, 2- and 3-Year Follow-Up (fup 1 to fup 3) for Psychoanalytic (PA; $n = 35$) and Psychodynamic (PD; $n = 31$) Therapy

			ES (within)				ES (between)				CS (%)			
		N	Post	fup 1	fup 2	fup 3	Post	fup 1	fup 2	fup 3	Post	fup 1	fup 2	fup 3
BDI	PA	35	2.4	2.3	2.3	2.6	0.2	0.3	0.3	0.7	85.7	85.7	85.7	91.4
	PD	31	2.1	2.0	2.0	2.0					77.4	77.4	74.2	77.4
SCL-90-GSI	PA	35	1.6	1.5	1.6	1.8	0.1	0.3	0.2	0.6	57.1	54.3	65.7	68.6
	PD	31	1.3	1.0	1.1	1.1					58.1	41.9	48.4	48.4
IIP	PA	35	1.6	1.7	1.7	2.0	0.2	0.5	0.4	0.4	51.4	57.1	57.1	65.7
	PD	31	1.4	1.0	1.2	1.6					48.4	38.7	54.8	64.5
F-SOZU	PA	33	0.6	0.7	0.7	0.8	0.2	0.5	0.5	0.6	12.1	9.1	12.1	9.1
	PD	30	0.4	0.3	0.3	0.2					3.3	6.7	3.3	3.3
SPC	PA	35	1.9	2.2			0.6	0.6			57.1	68.6		
	PD	31	1.3	1.5							32.3	38.7		
INTREX	PA	33	1.2	1.3	1.4	1.5	0.3	0.3	0.4	0.5	no norm data available			
intr pos	PD	30	0.6	0.8	0.6	0.7								

Note. See Table 2a for abbreviations.

time by including a 3-year follow-up period in order to examine treatment effects as distinguished from the natural course of the disorder (Roth & Fonagy, 2005).

Symptoms

We found that both treatments were highly effective in alleviating symptoms over time in terms of both depressive and global psychiatric symptom measures, including our measure of self-rated remission of depression. Still, psychoanalytic therapy proved more effective than psychodynamic therapy at the 3-year follow-up, and our first hypothesis was thus supported.

The psychoanalytic treatment in our study produced larger effects than long-term treatments examined in previous research. Leichsenring and Rabung's (2008) meta-analysis and de Maat and colleagues' (2009) review of long-term psychotherapy studies reported smaller ESs for depressive and general psychiatric symptoms compared to our effects. The psychoanalytic treatment in our study also produced larger effects than the treatment investigated by Knekt and colleagues (2008; 2011), and produced larger effects than the treatment examined by Grande and colleagues (2006). The higher effectiveness in our study can be attributed to the diagnostically homogenous groups, leading to higher effect sizes because of smaller variance, in contrast to the heterogeneous groups of these studies.

Based on psychoanalytic theory (e.g., Wallerstein 1986), one would not expect a substantial difference in outcome between psychoanalytic and psychodynamic therapy on a symptomatic level. Most previous studies support this hypothesis. The findings of the Psychotherapy Research Project of the Menninger Foundation (Wallerstein, 1986) supported this hypothesis, as observer-rated symptom measures revealed no significant difference between psychoanalysis and psychotherapy. The German Psychoanalytical Association study (Leuzinger-Bohleber, Rüger, Stuhr, & Beutel, 2002; Leuzinger-Bohleber et al., 2003) found no significant differences between psychoanalytic and psychodynamic patients in self-rated general well-being that included psychiatric symptoms. Grande and colleagues (2006) reported no significant differences in effect sizes for both treatment modalities on a self-rated symptom measure, and Jakobsen and

colleagues (2007) replicated these findings with data from four German studies on the effectiveness of psychoanalytic and psychodynamic therapies. But the results are still inconsistent. Weber, Bachrach, & Solomon (1985a; 1985b), in the Columbia Psychoanalytic Center Research Project, reported a significant difference between psychoanalysis and psychotherapy in observer-rated therapeutic benefit, an overall measure of symptomatic change. Sandell and colleagues (2000) found a significant difference in the growth curve analysis of the symptom measure they used in the comparison of psychoanalysis and psychotherapy, but this finding was not replicated in a subsample (Falkenström, Grant, Broberg, & Sandell, 2007).

Personality Functioning

We found that both treatments were highly effective in producing change in personality functioning, though psychoanalytic therapy was superior to psychodynamic therapy in terms of psychological capacities measured at 1-year follow-up and positive self-appraisal at 3-year follow-up. In addition, our findings suggest that the benefits of psychoanalytic therapy but not psychodynamic therapy continue to increase after termination in positive self-appraisal.

It is difficult to discuss our results in the context of the existing empirical evidence for long-term psychotherapy regarding personality change, in part because previous studies differ in salient study parameters (treatments, follow-up intervals, and measures). However, our findings build on the results of the Rudolf, Manz, and Oeri study reported by Fonagy and colleagues (2002) who found that psychoanalytic therapy produced more marked improvements in symptoms and structural personality functioning compared to psychodynamic therapy. Our findings further build on the conclusions by Leichsenring and Rabung (2008) who found increases in ESs for personality functioning from post-treatment to follow-up for long-term psychotherapies, indicating that the change processes are still going on after termination of treatment (see also Rice & Greenberg, 1984). But divergent results came from the studies of the Columbia Psychoanalytic Center Research Project (Weber et al., 1985a; Weber et al., 1985b) where ego strength was assessed to measure improvement over and above symptoms. Pre/post effect sizes were inconsistent: large for both groups in study I (Weber et al., 1985a) and large for psychoanalysis and small for psychotherapy in study II (Weber et al., 1985b). Wallerstein (1986) summarized the results of the Psychotherapy Research Project, indicating that "the psychotherapy cases did as well as the psychoanalytic ones." The authors of the German Psychoanalytical Association study (Leuzinger-Bohleber et al., 2002, 2003) reported significant differences between psychoanalysis and psychotherapy in the expert-rated dimensions of self-reflection, creativity, and ability to work, but surprisingly not in the relationship dimension. In the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (Sandell et al., 2000), the development of social functioning was virtually the same whether a patient had been in psychoanalysis or psychotherapy; mean improvement was almost exactly equal in both groups, and within-group effect sizes were small for both groups.

Personality functioning is hypothesized to play a contributing role in the development of depression (Akiskal, Hirschfeld, & Yerevanian, 1983; Kendler, Gatz, Gardner, & Pedersen, 2006; Krueger, 2005; Morey et al., 2007), and impairment in personality functioning, beyond its significance for social adjustment and life satisfaction, has been shown to predispose individuals for symptom persistence, relapse, and recurrence (Fava et al., 2007). Therefore, improvement in personality functioning may function as a mediator for the stability of symptom change, and Grande and colleagues (2009) indeed found that personality change at post-treatment was a predictor of patient self-reported symptom change at

3-year follow-up. While we did not conduct any mediational analyses, our results showing significant improvements in both symptoms and personality functioning in the psychoanalytic group lend some support to the mediation hypothesis.

Social Relations

We found that both treatments produced equally large improvements in interpersonal problems measured at 3-year follow-up, while psychoanalytic therapy was superior to psychodynamic therapy in terms of perceived social support. So our results are consistent with Leichsenring and Rabung's (2008) conclusion that long-term psychotherapies are associated with improvements in social functioning and with the findings of Grande and colleagues (2006) and Sandell and colleagues (2000) who found no significant differences between psychoanalysis and psychodynamic therapy for interpersonal problems after termination of treatment. The non-significant difference between psychoanalytic and psychodynamic therapy in change in interpersonal problems is counter-intuitive because according to psychoanalytic theory the more conflict-oriented technique of psychoanalytic therapy is expected to reduce interpersonal problems significantly better than psychodynamic therapy. We believe that psychoanalytic treatment produced significantly larger improvements in perceived social support, but not in interpersonal problems, because psychoanalytic therapy targets the depressed patients' interest in their environment and their more unprejudiced exploration of their environment (McGlashan & Miller 1982), thus enabling them to profit more extensively from their social network.

Although earlier comparative studies on the effectiveness of psychoanalysis are methodologically seriously flawed (Bachrach, Galatzer-Levy, Skolnikoff, & Waldon, 1991), there still is a considerable body of knowledge reported in the psychotherapy research supporting that clear or consistent differences on several dimensions of outcome among the therapies are lacking. We explain this discrepancy in our findings with two methodological reasons: (1) our randomized allocation to the treatment arms, that ensures comparability of patients in different treatment conditions, and (2) the diagnostic homogeneity of our patients, that enhances the power of the design to detect differences in effectiveness that exist. The obvious difference in treatment dose (PA = mean of 234 sessions in 39 months, PD = mean of 88 sessions in 34 months), which is an essential parameter of the treatment packages under scrutiny, may account for the differences across all dimensions of outcome.

Strengths and Limitations

We will discuss the strengths and limitations of our study in terms of the pragmatic-explanatory continuum indicator summary (PRECIS; Thorpe et al., 2009) to avoid simplistic labelling. According to PRECIS criteria, our study shares some characteristics with effectiveness studies (pragmatic studies which are more naturalistic and maximize external validity) and some characteristics with efficacy studies (explanatory studies which maximize internal validity) and thus falls in the middle of the effectiveness-efficacy continuum. The following features of the study were strengths in terms of external validity: Participants were enrolled under the routine conditions of a university outpatient clinic; they were "real-world" patients without restrictive selection criteria; the treatments were applied by "real-world" therapists; no treatment manuals were used; no special strategies like continuous supervision to maintain or improve adherence were used; and the patients were followed up for three years. At the same time, the randomized allocation to groups and the blindness of investigators to treatment modality served to increase internal validity. We consider our study to be in the middle of the pragmatic-explanatory continuum, having significant

scientific strengths as well as the potential capacity to inform healthcare decision-making regarding clinical practice.

Our study has several limitations as well, including the small sample sizes and the lack of a Structured Clinical Interview for *DSM-IV* (SCID-I and SCID-II) assessment of primary diagnoses and co-morbid disorders due to limited resources. The lack of SCID diagnoses may have caused an underestimation of Axis I and Axis II co-morbidity, thus limiting the comparability with other significant studies. We tried to compensate for this by applying the International Diagnostic Checklists for ICD-10 and *DSM-IV* through two experienced psychiatrists who decided consensually. Another limitation is the lack of a low-intensity treatment group, like treatment as usual, to control for the natural course of the disorder because depressive episodes in some patients are self-limited and typically remit in six to eight months (Berger and van Calker, 2004; Ustün, Ayuso-Mateos, Chatterji, Mathers, & Murray, 2004; Wittchen, 1988). On the other hand, patients were followed up three years after termination of treatment and would have relapsed if the treatments had not changed the natural course of the disorder.

In terms of internal validity, the differing dose of the treatments (i.e., the number of sessions) can be considered a confound. We believe that each treatment has a different underlying working model that needs a specified time frame, with a stipulated number of sessions and specific interventions in order to initiate a specific process. Moreover, we deliberately wanted to investigate treatment packages with their prototypical doses (according to the German Psychotherapy Guidelines) to inform practitioners about their everyday practice, which may limit the generalizability of the results, for example, by excluding four sessions per week psychoanalysis.

Another limitation and threat to internal validity (though a strength in terms of external validity) was the lack of treatment manuals and the absence of a manual-guided adherence measure. However, treatment fidelity was assessed, and the difference between the two treatments was confirmed by our adherence measures. Different investigators performed assessments at pre-treatment, post-treatment, and 1-year follow-up, with the advantage to blind them for treatment modality but the disadvantage of enhancing the variance; we tried to compensate for error variance by a joint training and by recalibration sessions. Besides, internal validity of the study would have profited by a random assignment of participants to investigators, blinded to measurement points.

Implications of the Study

This study adds some empirical evidence for the effectiveness of long-term psychotherapies, since both treatments in the study produced significant improvement from pre-treatment to post-treatment and follow-up. At the same time, we found psychoanalytic therapy to be superior to psychodynamic therapy in terms of symptoms, personality functioning, and (to some extent) social relations, demonstrating the full range of its benefits three years after termination of treatment. But these findings are far from being robust until they are replicated by independent study groups that can eliminate some of the threats to internal validity of our study. Although the research question of the differential outcome of long-term psychotherapy is far from being settled, subsequent studies should integrate a process-outcome approach to advance recent questions of psychotherapy research: What is the shape of change over time (trajectories of change)? Under what conditions does change occur (moderators of change)? How is change occurring (mediators of change)? (Laurenceau, Hayes, & Feldman, 2007). But although more research questions are open than answered, the results of our study lend some support to the effectiveness of long-term treatment generally and psychoanalytic treatment specifically, which has implications for healthcare policy regarding clinical practice.

REFERENCES

Ablon, J. S., & Jones, E. E. (2005). On psychoanalytic process. *Journal of the American Psychoanalytic Association, 53*, 541-568.

Akiskal, H. S., Hirschfeld, R. M., & Yerevanian, B. I. (1983). The relationship of personality to affective disorders. *Archives of General Psychiatry, 40,* 801-810.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) (*DSM-IV*). Washington, DC: Author.

Bachrach, H. M., Galatzer-Levy, R., Skolnikoff, A., & Waldon, S., Jr. (1991). On the efficacy of psychoanalysis. *Journal of the American Psychoanalytic Association, 39*, 871-916.

Beenen, F., & Stoker, J. (1996). *Periodical rating scale for psychoanalytic treatment*. Amsterdam: Netherlands Psychoanalytic Institute.

Benjamin, L. S. (1974). Structural analysis of social behavior. *Psychological Review, 81*, 392-425.

Benjamin, L. S. (1983). *The INTREX user's manual* (Parts I and II). Madison, WI: Intrex Interpersonal Institute.

Berger, M., & van Calker, D. (2004). Affektive Stoerungen [Affective disorders]. In M. Berger (Ed.), *Psychische Erkrankungen. Klinik und Therapie* [Psychic disorders. Clinical appearance and therapy]. Munich: Urban & Fischer.

Boyce, P., Parker, G., Barnett, B., Cooney, M., & Smith. F. (1991). Personality as a vulnerability factor to depression. *British Journal of Psychiatry*, *159*, 106-114.

Bush, M., & Meehan, W. (2011). Should supportive measures and relational variables be considered a part of psychoanalytic technique? Some empirical considerations. *International Journal of Psychoanalysis*, *92*, 377-399.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cuijpers, P., van Straten, A., Schuurmans, J., van Oppen, P., Hollon, S. D., & Andersson, G. (2010). Psychotherapy for chronic major depression and dysthymia: A meta-analysis *Clinical Psychology Review, 30*, 51-62.

De Maat, S., de Jonghe, F., Schoevers, R., & Dekker, J. (2009). The effectiveness of long-term psychoanalytic therapy: A systematic review of empirical studies. *Harvard Review of Psychiatry*, *17*, 1-23.

DeWitt, K. N., Milbrath, C., & Wallerstein, R. S. (1999). Scales of psychological capacities: Support for a measure of structural change. *Psychoanalysis and Contemporary Thought, 22*, 453-80.

DGPPN, BÄK, KBV, AWMF, AkdÄ, BPtK, et al. (2009). *S3-Leitlinie/nationale versorgungsleitlinie unipolare depression-langfassung* [S 3-guideline/national healthcare guideline "unipolar depression"]. Berlin: DGPPN, ÄZQ, AWMF.

Driessen, E., Cuijpers, P., de Maat, S. C., Abbass, A. A., de Jonghe, F., & Dekker, J. J.. (2010). The efficacy of short-term psychodynamic psychotherapy for depression: A meta-analysis. *Clinical Psychology Review, 30*, 25-36.

Dunkel, D., Antretter, E., Fröhlich-Walser, S., & Haring, C. (2005). Evaluation der Kurzform des Fragebogens zur Sozialen Unterstützung (SOZU-K-22) in klinischen und nichtklinischen stichproben [Evaluation of the short-form Social Support Questionnaire (SOZU-K-22) in clinical and non-clinical samples]. *Psychotherapy, Psychosomatics, Medical Psychology, 55,* 266-277.

Dunner, D. L. (2001). Acute and maintenance treatment of chronic depression. *Journal of Clinical Psychiatry*, *62*(Suppl. 6), 10-16.

Eaton, W. W. (2008). Population-based study of first onset and chronicity in major depressive disorder. *Archives of General Psychiatry*, *65*, 513-520.

Falkenström, F., Grant, J., Broberg, J., & Sandell, R. (2007). Self-analysis and post-termination improvement after psychoanalysis and long-term psychotherapy. *Journal of the American Psychoanalytic Association*, *55*, 629-674.

Fava, G. A., Ruini, C., & Belaise, C. (2007). The concept of recovery in major depression. *Psychological Medicine, 37*, 307-317.

Fonagy, P., Kaechele, H., Krause, R., Jones, E., Perron, R., Clarkin, J., et al. (2002). *An open door review of outcome studies in psychoanalysis*

(2nd ed.). London: International Psychoanalytic Association.

Fonagy, P., & Kaechele, H. (2009). Psychoanalysis and other long-term dynamic psychotherapies. In M.-G. Gelder, J. J. Lopez-Ibor, and N. Andreasen (Eds.), *New Oxford textbook of psychiatry*, Vol. 2 (pp. 1337–1349). Oxford: Oxford University Press.

Frank, E., Kupfer, D. J., Perel, J. M., Cornes, C., Jarrett, D. B., Mallinger, A. G., et al. (1990). Three-year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry, 47*, 1093-1099.

Franke, G. (1995). *Die symptom-checkliste von Derogatis. Deutsche version. Manual* [The symptom-checklist of Derogatis. German version. Manual]. Goettingen: Beltz.

Fydrich, T., Geyer, M., Hessel, A., Sommer, G., & Brähler, E. (1999). Fragebogen zur sozialen unterstützung (F-SozU): Normierung an einer Stichprobe [Social Support Questionnaire (F-SozU): Norms of a representative sample]. *Diagnostica*, *45*, 212-216.

Gabbard, G. O., & Westen, D. (2003). Rethinking therapeutic action. *International Journal of Psychoanalysis*, *84*, 823-841.

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., Dilg, R., Jakobsen, T., Keller, W., Krawietz, B., Langer, M., et al. (2009). Structural change as a predictor of long-term follow-up outcome. *Psychotherapy Research*, *19*, 344-357.

Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (1994). *Beck Depressions-Inventar* (BDI) [Beck Depression Inventory]. Bern: Huber.

Henry, W. P., Schacht, T. E., & Strupp, H. H. (1990). Patient and therapist introject, interpersonal process, and differential psychotherapy outcome. *Journal of Consulting and Clinical Psychology, 58*, 768-774.

Hiller, W., Dichtl, G., Hecht, H., Hundt, W., Mombour, W., & von Zerssen, D. (1994). Evaluating the new ICD-10 categories of depressive episode and recurrent depressive disorder. *Journal of Affective Disorders, 31*, 49-60.

Hiller, W., Zaudig, M., & Mombour, W. (1995). *IDCL – Internationale diagnosen checklisten für ICD-10 und DSM-IV (Manual und 32 checklisten nach ICD-10 als bestandteil des gesamtpaketes der ICD-10-checklisten der WHO)* [International diagnostic checklists for ICD-10 and DSM-IV (Manual and 32 checklists as part of the WHO ICD-10-checklists package)]. Bern: Huber.

Horowitz, L. M., Strauß, B., & Kordy, H. (2000). *Inventar zur Erfassung Interpersoneller Probleme* [Inventory of Interpersonal Problems]. Goettingen: Beltz.

Huber, D., Brandl, T., & Klug, G. (2004). The Scales of Psychological Capacities (SPC): Measuring beyond symptoms. *Psychotherapy Research*, *14*, 89-106.

Huber, D., Henrich, G., & Klug, G. (2005). The Scales of Psychological Capacities: Measuring change in psychic structure. *Psychotherapy Research*, *15*, 445-456.

Huber, D., Henrich, G., & Klug, G. (2007). The Inventory of Interpersonal Problems (IIP): Sensitivity to change. *Psychotherapy Research*, *17*, 474-481.

Huber, D., Klug, G., & Wallerstein, R. S. (2006). *Skalen psychischer kompetenzen (SPK). Ein messinstrument für therapeutische veränderung in der psychischen struktur. Manual und Interviewleitfaden* [The Scales of Psychological Capacities (SPC). A measure for therapeutic change in psychic structure. Manual and interview guidelines]. Stuttgart, Germany: Kohlhammer.

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.

Jakobsen, T., Rudolf, G., Brockmann, J., Eckert, J., Huber, D., Klug, G., et al. (2007). Ergebnisse analytischer Langzeitpsychotherapien bei spezifischen psychischen Störungen: Verbesserungen in der Symptomatik und in den interpersonellen Beziehungen [Results of psychoanalytic long-term therapy in specific diagnostic groups: Improvement in symptoms and interpersonal relationships]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, *53*, 87-110.

Jones, E. E. (2000). *Therapeutic action. A guide to psychoanalytic therapy* Northvale, NJ: Aronson.

Judd, L. L. (1991). The clinical course of unipolar major depressive disorders. *Archives of General Psychiatry*, *54*, 989-991.

Kaechele, H. (2010). Distinguishing psychoanalysis from psychotherapy. *International Journal of Psychoanalysis*, *91*, 35-43.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, *3*, 1-27.

Keller, M. B., Hirschfeld, R. M. A., & Hanks, D. (1997). Double depression: A distinctive subtype of unipolar depression. *Journal of Affective Disorders, 45*, 65-73.

Keller, M. B., Lavori, P. W., Mueller, T. I., Endicott, J., Coryell, W., Hirschfeld, R. M., et al. (1992). Time to recovery, chronicity, and levels of psychopathology in major depression: A 5-year prospective follow-up of 431 subjects. *Archives of General Psychiatry*, *49*, 809-816.

Keller, M. B., & Shapiro, R. W. (1981). Major depressive disoerder: Initial results from a one-year prospective naturalistic follow-up study. *Journal of Nervous and Mental Disease, 169*, 761-768.

Kendler, K. S., Gatz, M., Gardner, C. O., & Pedersen, N. L. (2006). Personality and major depression: A Swedish longitudinal, population-based twin study. *Archives of General Psychiatry, 63*, 1113-1120.

Klug, G., Henrich, G., Kaechele, H., Sandell, R., & Huber, D. (2008). Die Therapeutenvariable: Immer noch ein dunkler Kontinent? [The therapist variable: Is it still an unknown factor?]. *Psychotherapeutics, 53*, 83-91.

Klug, G., & Huber, D. (2009). Psychic structure: Exploring an empirically still unknown territory. *Journal of the American Psychoanalytic Association*, *57*, 149-173.

Knekt, P., Lindfors, O., Härkänen, T., Välikoski, M., Virtala, E., Laaksonen, M. A., et al. (2008). Randomized trial on the effectiveness of long- and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychological Medicine*, *38*, 689-703.

Knekt, P., Lindfors, O., Laaksonen, M. A., Renlund, C., Haaramo, P, Härkänen, T., et al. (2011). Quasi-experimental study on the effectiveness of psychoanalysis, long-term and short-term psychotherapy on psychiatric symptoms, work ability and functional capacity during a 5-year follow-up. *Journal of Affective Disorders, 132*, 37-47.

Koppers, D., Peen, J., Niekerken, S., Van, R., & Dekker, J. (2011). Prevalence and risk factors for recurrence of depression five years after short-term psychodynamic treatment. *Journal of Affective Disorders, 134*, 468-472.

Krueger, R. F. (1999). Personality traits in late adolescence predict mental disorders in early adulthood: A prospective-epidemiological study. *Journal of Personality, 67*, 39-65.

Krueger, R. F. (2005). Continuity of axes I and II: Toward a unified model of personality, personality disorders, and clinical disorders. *Journal of Personality Disorders, 19*, 233-261.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. Lambert (Ed.), *Handbook of psychotherapy and behavior change*. Hoboken, NJ: Wiley.

Laurenceau, J.-P., Hayes, A. M., & Feldman, C. G. (2007). Some methodological and statistical issues in the study of change processes in psychotherapy. *Clinical Psychology Review, 27*, 682-695.

Leichsenring, F., & Rabung, S. (2008). Effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *Journal of the American Medical Association, 300*, 551-1565.

Leichsenring, F., & Rabung, S. (2011). Long-term psychodynamic psychotherapy in complex mental disorders: Update of a meta-analysis. *British Journal of Psychiatry*, *199*, 15-22.

Leuzinger-Bohleber, M., Rüger, B., Stuhr, U., & Beutel, M. (2002). "*Forschen und Heilen*" *in der Psychoanalyse* ["Researching and healing" in psychoanalysis]. Stuttgart: Kohlhammer.

Leuzinger-Bohleber, M., Stuhr, U., Rueger, B., & Beutel, M. (2003). How to study the "quality of psychoanalytic treatments" and their long-term effects on patients' well-being: A epresentative, multi-perspective follow-up study. *International Journal of Psychoanalysis, 84*, 263-290

Luyten, P., & Blatt, S. J. (2011). Psychodynamic approaches to depression: Whither shall we go? *Psychiatry*, *74*, 1-3.

McGlashan, T. A., & Miller, G. H. (1082). The goals of psychoanalysis and psychoanalytic psychotherapy. *Archives of General Psychiatry*, *39*, 377-388.

Morey, L. C., Hopwood, C. J., Gunderson, J. G., Skodol, A. E., Shea, M. T., Yes, S., et al. (2007). Comparison of alternative models for personality disorders. *Psychological Medicine*, *37*, 983-994.

Rice, L. N., & Greenberg, L. S. (1984). *Patterns of change*. New York: Guilford.

Roth, A., & Fonagy, P. (2005). *What works for whom?* New York: Guilford.

Rueger, U., Dahm, A., & Kallinke, D. (2005). *Faber-Haarstrick: Kommentar psychotherapierichtlinien* [The Faber-Haarstrick commentary on psychotherapy guidelines]. Munich: Urban & Fischer.

Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J., & Schubert, J. (2000). Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPP). *International Journal of Psychoanalysis, 81*, 921-942.

Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., et al. (1992). Course of depressive symptoms over follow-up: Findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Archives of General Psychiatry*, *49*, 782-787.

Sommer, G., & Fydrich, T. (1991). Entwicklung und ueberpruefung eines fragebogens zur sozialen unterstuetzung [Development and evaluation of a questionnaire for social support]. *Diagnostica*, *37*, 160-178.

Taylor, D. J., Walters, H. M., Vittengl, J. R., Krebaum, S., & Jarrett, R. B. (2010). Which depressive symptoms remain after response to cognitive therapy of depression and predict relapse and recurrence? *Journal of Affective Disorders, 123*, 181-187.

Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G., et al. (2009). A pragmatic-explanatory continuum indicator summary (PRECIS): A tool to help trial designers. *Journal of Clinical Epidemiology*, *62*, 464-475.

Tress, W. (1993). INTREX-Kurzform-Fragebogen: Formulare, itemformulierungen, und fragestellungen [INTREX questionnaire, short form: Forms, item formulation, and formulation of questions]. In W. Tress (Ed.), *SASB: Die strukturale analyse sozialen verhaltens* [SASB: Structural Analysis of Social Behavior] (pp. 259-263). Heidelberg: Asanger.

Ustün, T. B., Ayuso-Mateos, J. L., Chatterji, S., Mathers, C., & Murray, C. J. (2004). Global burden of depressive disorders in the year 2000. *British Journal of Psychiatry*, *184*, 386-392.

Vittengl, J. R., Clark, L. A., & Jarrett, R. B. (2009). Deterioration in psychosocial functioning predicts relapse/recurrence after cognitive therapy for depression. *Journal of Affective Disorders*, *112*, 135-143.

Wallerstein, R. S. (1986). *Forty-two lives in treatment*. New York: Guilford.

Weber, J., Bachrach, H., & Solomon, M. (1985a). Factors associated with the outcome of psychoanalysis: Report of the Columbia Psychoanalytic Center Research Project (II). *International Review of Psychoanalysis*, *12*, 127-141.

Weber, J., Bachrach, H., & Solomon, M. (1985b). Factors associated with the outcome of psychoanalysis: Report of the Columbia Center Research Project (III). *International Review of Psychoanalysis*, *12*, 251-262.

World Health Organization (WHO). (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva, Switzerland: World Health Organization.

Williams, J. W., Noël, P. H., Cordes, J. A., Ramirez, G., & Pignone, M. (2002). Is this patient clinically depressed? *Journal of the American Medical Association*, *287*, 1160-1170.

Wittchen, H.-U. (1988). Zum spontanverlauf unbehandelter fälle mit angststörungen bzw. depressionen [Natural course of untreated cases of anxiety disorders respectively depressions]. In H.-U. Wittchen and D. von Zeerssen (Eds.), *Verläufe behandelter und unbehandelter depressionen und angststörungen* [Courses of treated and untreated depressions and anxiety disorders] (pp. 252–284). New York: Springer.