

Effectiveness of Integrative Imagery- and Trance-Based Psychodynamic Therapies: Guided Imagery Psychotherapy and Hypnotherapy

Christian Sell and Heidi Möller
University of Kassel

Svenja Taubner
Heidelberg University Hospital

In a naturalistic study, we examined the effectiveness of two integrative psychodynamic psychotherapies—Guided Imagery Psychotherapy (GIP) and Hypnotherapy (HY) – in an outpatient setting. A mixed sample of patients was assessed at the beginning of therapy ($N = 300$) and after 30 months ($N = 97$), using the IIP, the IPO-16, the FLZ^M, and the PMS as well as health economic measures. The BSI was employed at the beginning of therapy and every 6 months after that. Therapists rated CGI-S and GAF at the beginning and end of therapy. There was no control group. Improvements were found on all measures for GIP and for all but 1 measure for HY. Effect sizes ranged from $d = .12$ to $d = 1.76$. The percentage of patients who changed reliably ranged from 18% to 82.5%. We statistically controlled for the number of treatment sessions and whether the treatment had been completed within 30 months. For the BSI, improvement was found within the first 6 months of treatment and again within the following 18 to 24 months—indicating a regressive but steady alleviation of symptom distress over the course of longer treatments. In conclusion, under naturalistic conditions, GIP and HY have been shown to be effective with regards to different outcome modalities.

Keywords: outcome research, mental health services research, guided imagery, hypnosis, assimilative psychodynamic therapy

There is a continuously growing research interest in understanding the role of imagery in mental processes and in evaluating ways in which imagery could be used in fields of applied psychology (Forrester, 2000; Röckelein, 2004). In psychotherapy, several approaches

utilize imagery for different functions in the therapeutic process (e.g., Brewin et al., 2009; Edwards, 2007; Holmes & Mathews, 2010). Typically, these therapies can be described as forms of assimilative psychotherapy integration (Arkowitz, 1997). This refers to methods of psychotherapy that operate within the framework of one established school of therapy but expand the established canon of that school's treatment technique by integrating interventions from other approaches. This strategy of integration is meant to combine “the advantage of a single, coherent theoretical system with the flexibility of a broader range of technical interventions from multiple systems” (Norcross, 2005, p. 10). Examples include Assimilative Psychodynamic Psychotherapy (Stricker & Gold, 2005), Cognitive–Behavioral Assimilative Integration (Castonguay, Newman, Borkovec, Holtforth, & Maramba, 2005), integration in humanistic psychotherapies (West, 2000), and Gestalt Reminiscence Therapy (O’Leary & Barry, 2006). With regards to integrating imag-

Christian Sell and Heidi Möller, Department of Psychology, University of Kassel; Svenja Taubner, Institute for Psychosocial Prevention, Heidelberg University Hospital.

For the present study the University of Kassel received third-party funding from the *Austrian Society for Applied Depth-Psychology and General Psychotherapy* (ÖGATAP). We hereby confirm that ÖGATAP had no influence on the study design, the analyses, the interpretation of the results, or the decision to publish. The study design was approved of by the ethics committee of the University of Kassel, Germany as well as by the local medical ethics committees in Vienna and Innsbruck, Austria.

Correspondence concerning this article should be addressed to Christian Sell, Department of Psychology, University of Kassel, Holländische Straße 36-38, 34127 Kassel, Germany. E-mail: csell@uni-kassel.de

ery into an established system of psychotherapy, there are several important developments within the cognitive-behavioral (CBT) tradition. Zinbarg, Craske, and Barlow (2006) have introduced a manual for treating Generalized Anxiety Disorder in which imagery exposure is central. Imagery also plays a prominent role in schema-focused therapies (Young, Klosko, & Weishaar, 2003), such as Imagery Rescripting and Reprocessing Therapy (Arntz, 2011, 2012; Schmucker & Dancu, 1999), as well as in Compassion Focused Therapy (Gilbert, 2010; Hackmann, 2005). In the psychodynamic tradition, Guided Imagery Psychotherapy (GIP) and Hypnopsychotherapy (HY), two smaller but long established therapeutic schools, have also developed concepts and techniques for incorporating imagery interventions, such as guided imagery or hypnotic trances, into their framework.

GIP is built upon Leuner's (1984) work on guided affective imagery. It is defined by Pichler (2011) as an integrative psychodynamically oriented method of psychotherapy during which imagery plays a central role in the therapeutic situation. The therapist is meant to work according to the basic principles of psychodynamic therapy—Leuner (1984) most prominently references Freudian drive theory and object relations theory—but will also regularly have sessions that involve guided imagery. The specific role of imagery (see the Treatment Procedures section for more details) is related to the idea of change in GIP: the specific features that a patient imagines a certain motif to have, such as a meadow or a mountain, are understood as a symbol for some of his or her inner object representations. The imaginations themselves as well as discussing their contents with the therapist are meant to lead the patient to a better understanding and subsequently to a modification of his or her “inner world.” In accordance with psychoanalytic theory, a change of the patient's inner world should in turn lessen the need for rigid defense mechanisms and thereby also alleviate psychopathological symptoms and interpersonal problems.

Bahrke and Nohr (2013) list GIP as an officially recognized therapeutic method of its own in Austria with approximately 480 practicing therapists, in Switzerland with 57 therapists, in Russia, Ukraine, Belarus, and Kazakhstan with a total of 285 therapists, and as an “additional method” in psychodynamic therapies in Ger-

many with 253 active therapists. Furthermore, as described by Hall, Hall, Stradling, and Young (2006), guided imagery as a set of therapeutic techniques is widely used in other European countries as well—often as part of or as an extension to different forms of psychodynamically oriented inpatient and outpatient therapies. There are professional societies for GIP in Austria, Switzerland, Germany, the Netherlands, the Czech Republic, Slovakia, Russia/Ukraine/Belarus/Kazakhstan, and Sweden.

HY is described by Kanitschar (2009) as an integrative psychodynamic method open to all known approaches and techniques of psychotherapeutic hypnosis. This entails the integration of hypnotherapeutic techniques that are not psychodynamic in their origin. Historically, HY has been based on the psychodynamic approach of hypnoanalysis as developed by Watkins and Fromm (Watkins, 1992). Brown and Fromm (1986, p. 196f) identify different psychodynamic models as relevant to hypnoanalysis: Whereas therapy with neurotic patients is primarily based upon Freudian drive theory, treatment of severe personality disorder emphasizes work on psychic structure as addressed by object relations theory and self-psychology. Fromm (1977) then used the terminology of ego psychology to merge these different psychoanalytic theories into an integrative conceptual framework for hypnosis-based psychodynamic therapies. HY in its present form subsequently developed through the integration of Ericksonian elements, as well as techniques from systemic and cognitive-behavioral hypnotherapy. Within an explicitly psychodynamic framework, the general idea of change in HY is to identify unconscious conflicts underlying a patient's pathology and offering corrective emotional experiences that correspond to these conflicts. Inducing hypnotic trances is hereby regarded as a means to loosening defense mechanisms and making unconscious material conscious (see Treatment Procedures section for more details).

Hypnopsychotherapy, as in the specific integration described above, is currently only officially recognized as a therapeutic method of its own in Austria with approximately 120 practicing therapists. There are, however, further training modules in HY (250–300 hours) available to medical doctors and psychotherapists in Germany (currently 175 practicing therapists) and

Switzerland (approximately 150 therapists). In addition to HY, the integration of hypnosis into psychotherapy is internationally widespread. The European Society of Hypnosis counts 3168 psychotherapists among its members, and national societies exist in most European countries. Other prominent forms of hypnotherapy include Neo-Ericksonian hypnotherapy (Erickson & Rossi, 1981) and cognitive-behavioral hypnotherapy (Robertson, 2012). These therapies differ considerably with regard to the types of hypnosis interventions they implement. In addition to the induction of corrective experiences through imagery, hypnosis is also often used for pain management, relaxation, or the exploration of thought patterns and traumatic memories.

Efficacy and effectiveness of psychodynamic psychotherapies in general are well established for a considerable number of disorders (Leichsenring & Rabung, 2011; Levy, Ehrental, Yeomans, & Caligor, 2014; Shedler, 2012), including depression, anxiety disorders, and personality disorders (Barber, Muran, McCarthy, & Keefe, 2013). Up until now, however, there has been no larger-scale efficacy research on outpatient GIP or HY specifically and very little research on the methods' effectiveness. Two prospective effectiveness studies on GIP have been published so far: In a German sample ($N = 66$), von Wietersheim, Wilke, Röser, and Meder (2003) found large effect sizes for depressive symptoms ($d = .84$) and life satisfaction ($d = .97$) as well as medium size effects for anxiety-related symptoms ($d = .49$) and somatic symptoms ($d = .44$). Sachsse, Imruck, and Bahrke (2016), again in a German sample ($N = 25$), found very large effects for symptom distress ($d = 1.41$) and large effects for interpersonal problems ($d = .81$) as well as for depression- and anxiety-related symptoms ($d = 1.14$). Furthermore, there is a considerable number of process and single case studies on GIP (e.g., Stigler & Pokorny, 2001; overview: Stigler & Pokorny, 2012) as well as some studies looking at GIP-treatment results for specific psychosomatic conditions or eating disorders. For patients with bulimia nervosa, for example, Esplen, Garfinkel, Olmsted, Gallop, and Kennedy (1998) found short-term GIP to be significantly more effective than a control condition where participants attended a weekly support-group meeting.

Most of the research available on clinical hypnosis in general focuses on the medical sector. Efficacy and effectiveness are considered to be established in pain management and in reducing stress associated with medical procedures (Hagl, 2013). Concerning the psychotherapeutic treatment of mental disorders, a meta-analysis by Revenstorf (2006) showed large effects for patients with anxiety disorders ($\bar{d} = .95$) and certain mild forms of addiction ($\bar{d} = .89$), as well as medium size effects for chronic pain disorders ($\bar{d} = .54$). In a meta-analysis from 2014, Hagl (2014) reported that recent research on hypnosis did not involve any clinical studies with hypnosis-based individual psychotherapy. HY specifically has not yet been subjected to systematic efficacy or effectiveness research at all.

In summary, there has already been some research on the therapeutic effectiveness of GIP as well as hypnosis in general but the body of evidence is far from conclusive. However, if these approaches to integrative psychodynamic therapy are not examined specifically, their unique contribution will not be appreciated by policymakers. Consequently, the goal of this study was a multimethod assessment of long-term therapeutic change in GIP and HY¹. Outcome criteria should involve a broad range of symptom-related and more general measures. Previous findings (Sell, Möller, & Taubner, 2016) suggest that imagery- and trance-based psychodynamic therapies are more effective for patients with higher levels of psychological mindedness; we included psychological mindedness as an outcome measure to assess whether prolonged therapeutic engagement with GIP or HY can also increase this capacity. Since outcome research for GIP and HY is at its early stages, we were at this point in time primarily interested in assessing both methods' effectiveness in general. The principles of naturalistic research, however, imply that we would have to ask the participating therapists to carry out their treatment as they normally would. We thereby allowed for different treatment lengths, session frequencies, and other variations to be present in the sample. Since it has long been argued that

¹ The original study design also envisaged an assessment of Autogenic Psychotherapy. Due to a poor sample size, however, this could not be realized.

there is a dose–response effect in psychotherapy (Hansen, Lambert, & Forman, 2002; Smit et al., 2012)—Orlinksy, Grawe, and Parks (1994) found a positive correlation between outcome and treatment length in 100 out of 156 studies—we aimed at controlling for number of sessions in all analyses.

As Stricker (2010) pointed out, effectiveness research on integrative therapies is particularly challenging because—while absolutely necessary—conducting it with too narrow a framework that restricts the flexibility of the therapists would be “contrary to the spirit of psychotherapy integration” (p. 403). Following the call by DeFife et al. (2015) for the implementation of prospective psychotherapy studies in clinical practice settings, we therefore took practice-based/naturalistic research to be the methodology of choice. The Austrian health care system seemed like a particularly well-suited place for this project. Even though, as described above, GIP and HY are currently employed within the health care delivery systems of several countries, Austria is unique among them because both GIP and HY are among the officially recognized scientific-psychotherapeutic methods in this country. Together both methods’ practitioners make up a total of 9% of state-licensed psychotherapists in Austria.

Method

Research Design and Sampling Procedures

This study is a prospective longitudinal study following a naturalistic one-group repeated measures design with multiple outcome measures. Participants were outpatients seeking psychotherapeutic help within the Austrian health care system. Measurements were taken at the beginning of treatment and every six months from then on, leading to a total of six measurement points over the course of 30 months (see Figure 1 for measurement points). There were no exact statistics available to us on how long these treatments commonly last. However, from consulting with senior GIP and HY therapists we expected the sample to consist of a mixture of treatment lengths, including short-term (<25 sessions) and long-term treatments (>45 sessions). We set the total duration of the study to 30 months as a compromise between feasibility and long-term perspective. We used a short

questionnaire every six months and did a full assessment at the beginning of therapy and after 30 months because our focus was meant to be on long-term effects, that is, how well is the patient 30 months after beginning treatment? The therapists rated their patients at the beginning and at the end of treatment. Having an individual posttreatment assessment for the patient-reported measures as well proved impractical because there was no reliable way of determining in time when a given treatment had been completed.

The inclusion criteria for patients were as follows: age between 18 and 65 years; neither psychotic disorders nor severe drug addiction nor severe neurological disorders should be the focus of psychopathology or treatment (according to the therapist’s clinical judgment); at least one mental disorder requiring treatment; sufficient German language proficiency to understand the questionnaires; informed consent.²

We carried out therapist and patient recruitment in cooperation with the Austrian Society for Applied Depth-Psychology and General Psychotherapy (Österreichische Gesellschaft für angewandte Tiefenpsychologie und allgemeine Psychotherapie, ÖGATAP), which has represented and trained GIP and HY in Austria since 1969. Recruitment is detailed in Figure 1. We contacted every GIP and HY therapist in independent practice who was an active member of ÖGATAP. During a recruitment period of one year, those therapists who had agreed to participate approached each of their new patients who met the inclusion criteria and asked them to participate in the study. The therapists handed their patients detailed information about the study, a consent form, the initial patient-questionnaire, and a postage-paid return envelope.

² In Austria, psychotherapy is financed through a mixture of private payment by the patient (with a supplementary grant of €21.80 per session through the statutory health insurance), private complementary insurances, and a fixed quota of therapy sessions paid for every year by the statutory health insurance in each of the Austrian federal states. The accessibility of psychotherapy is therefore to varying degrees determined by place of residence and socioeconomic status. We may reasonably assume that the sample will therefore be biased towards a higher socioeconomic status when compared to countries where psychotherapy is fully covered by the statutory health insurances—Germany in particular.

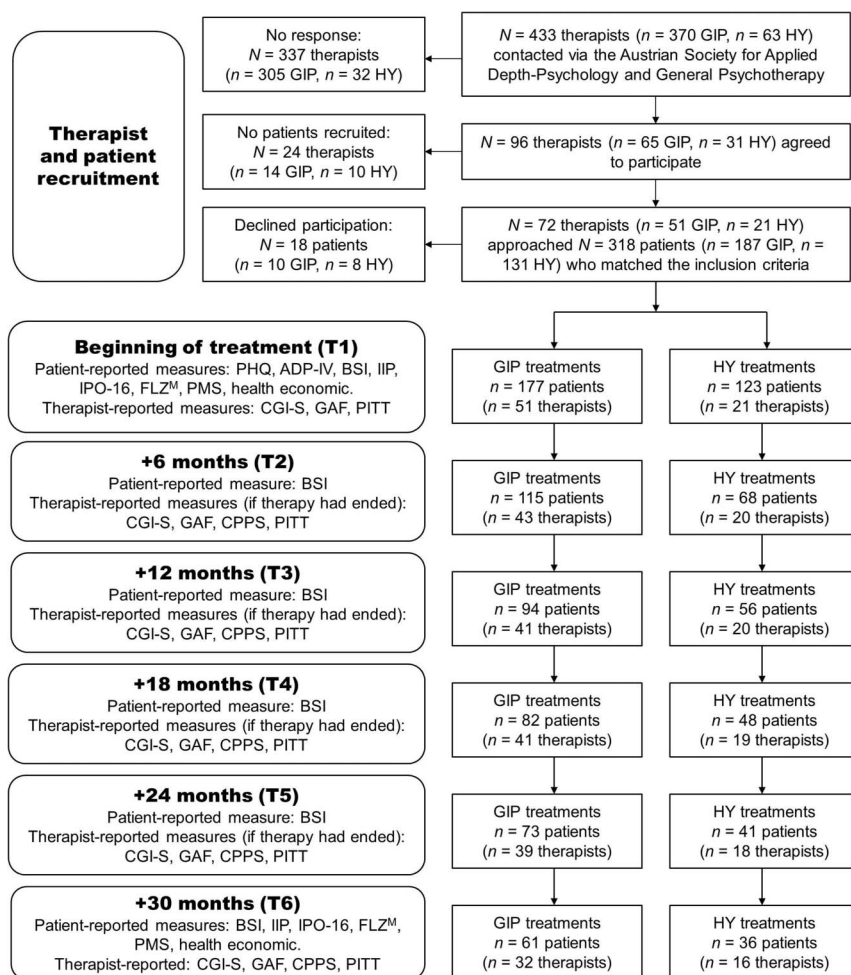


Figure 1. Therapist and patient flow chart, measurement points, and measures used. GIP = Guided Imagery Psychotherapy; HY = Hypnosispsychotherapy; PHQ = PRIME-MD Patient Health Questionnaire; ADP-IV = Assessment of *DSM-IV* personality disorders; BSI = Brief Symptom Inventory; IIP = Inventory of Interpersonal Problems; IPO-16 = Short form of the inventory of Personality Organization; FLV^M = Questions of Life Satisfaction; PMS = Psychological Mindfulness Scale; CGI-S = Clinical Global Impression—Severity; GAF = Global Assessment of Functions; CPPS = Comparative Psychotherapy Process Scale; PITT = Process Scale for Imagery- and Trance-Based Psychodynamic Therapies.

lope. We instructed the patients to mail all questionnaires directly to us, the research team—their respective therapists would not know whether a given patient agreed to participate in the study or not. Thenceforth all study-related communication with the patients was handled directly by us without involving the therapists. Whenever a study patient did not respond within six weeks after receiving a questionnaire, we sent a reminder together with another return

envelope. In order to preserve the confidentiality about a patient's decision whether or not to participate in the study, the therapists were asked to complete the therapist-questionnaires for every patient they had initially approached. Whenever we received therapist-questionnaires pertaining to patients who had chosen not to participate, we did not process the data and destroyed the physical copy of the questionnaire.

We made several attempts to recruit a control group of comparable patients treated only by a general practitioner (GP) and not by a psychotherapist. Unfortunately, these efforts were unsuccessful insofar as they did not yield enough patient data. We contacted GPs through the association of general practitioners in Austria (Hausärzterverband) as well as through several personal contacts in different provinces. Twelve GPs agreed to participate. We repeatedly tried to facilitate patient recruitment by providing the GPs with ready-to-use packages of information and questionnaires, offering a small financial incentive as well as individual feedback on the study's results. In the end, however, we received fewer than 10 initial patient-questionnaires and thus saw no other option but to drop the control group. The study is consequently limited to using comparable samples from the literature in order to evaluate the clinical significance of therapeutic change.

Participant Characteristics

Table 1 presents sample demographics as well as clinical information, separately for GIP and HY treatments. Patient recruitment yielded a total sample of $N = 300$ ($N_{\text{GIP}} = 177$, $N_{\text{HY}} = 123$) at the beginning of treatment (T1). Dropouts kept decreasing the sample size for the remaining measurement points (see Figure 1). The sample at T1 did not differ from the remaining sample at T6 on any of the preintervention measures for either GIP³ or HY⁴. Dropout was therefore not considered to be selective (see Discussion section for implications and possible predictors of dropouts).

Fifty-nine percent of the participants were on some form of psychiatric medication at T1 and 59.3% had some form of psychological therapy prior to this study. According to the PHQ, mood disorders were the most common diagnosis in the sample, followed by anxiety disorders and somatic symptom disorders. Twenty-eight percent of the patients had at least one personality disorder according to the ADP-IV. On average, patients had $M = 1.38$ ($SD = 1.33$, range: 0–5) axis I diagnoses and $M = .34$ ($SD = .76$, range: 0–4) axis II diagnoses. The majority of patients with a personality disorder (85.9%) had an axis I comorbidity.

The patients were treated by a total of 72 therapists (51 for GIP and 21 for HY).⁵ The patient-

therapist ratio was not significantly different between GIP and HY, $\chi^2(1) = 3.43$, $p = .06$, $r_\phi = .11$. The study therapists were predominantly seasoned practitioners with on average more than 15 years of professional experience. The majority of the therapists reported attending regular clinical supervision and more than a third of them were certified as supervisors themselves.

Measures

The full patient questionnaire was used at the beginning of treatment (T1) and after 30 months (T6), a brief patient questionnaire was used every 6 months in between (T2–T5). The therapists rated their patients at the beginning and end of the treatment. To assess adherence, the therapists indicated at the beginning of treatment which interventions and techniques they intended to use and at the end of treatment which interventions and techniques they had actually used (CPPS, PITT). Figure 1 summarizes the measures used on each measurement point.

³ Dropout analysis GIP: Age_{T1}: $t(236) = .93$, $p = .35$, $d = .08$; Gender_{T1}: $\chi^2(1) = 1.82$, $p = .18$, $r_\phi = .09$; BSI GSI_{T1}: $t(236) = .19$, $p = .85$, $d = .03$; IIP total_{T1}: $t(236) = .24$, $p = .81$, $d = .09$; IPO-16 total_{T1}: $t(236) = .72$, $p = .47$, $d = .06$; FLZ^M General_{T1}: $t(236) = .40$, $p = .69$, $d = .03$; FLZ^M Health_{T1}: $t(236) = .03$, $p = .98$, $d = .04$; PMS-D_{T1}: $t(236) = .96$, $p = .34$, $d = .13$; doctors' visits_{T1}: $t(236) = .11$, $p = .91$, $d = .02$; days of sick leave_{T1}: $t(236) = .30$, $p = .77$, $d = .06$; days in hospital_{T1}: $t(236) = .86$, $p = .39$, $d = .17$; CGI-S_{T1}: $t(236) = .81$, $p = .42$, $d = .07$; GAF_{T1}: $t(236) = .93$, $p = .35$, $d = .09$.

⁴ Dropout analysis HY: Age_{T1}: $t(157) = 1.09$, $p = .28$, $d = .23$; Gender_{T1}: $\chi^2(1) = .008$, $p = .93$, $r_\phi = .01$; BSI GSI_{T1}: $t(157) = 1.39$, $p = .17$, $d = .24$; IIP total_{T1}: $t(157) = 0.89$, $p = .37$, $d = .17$; IPO-16 total_{T1}: $t(157) = .40$, $p = .69$, $d = .06$; FLZ^M General_{T1}: $t(157) = .06$, $p = .95$, $d = .01$; FLZ^M Health_{T1}: $t(157) = .83$, $p = .41$, $d = .15$; PMS-D_{T1}: $t(157) = .90$, $p = .37$, $d = .16$; doctors' visits_{T1}: $t(157) = .12$, $p = .91$, $d = .02$; days of sick leave_{T1}: $t(157) = .19$, $p = .85$, $d = .04$; days in hospital_{T1}: $t(157) = .22$, $p = .82$, $d = .09$; CGI-S_{T1}: $t(157) = .46$, $p = .65$, $d = .09$; GAF_{T1}: $t(157) = .33$, $p = .74$, $d = .06$.

⁵ There are an unequal number of therapists for both methods because when planning the study we were unable to predict how many patients each study therapist would be able to recruit. To maximize the number of patients, we tried to include as many therapists as possible for both methods. The final numbers reflect that there are significantly more GIP therapists than HY therapists registered in Austria.

Table 1
Sample Demographics, Treatment Characteristics, and Clinical Information

| Characteristic | Guided Imagery (GIP) | Hypnotherapy (HY) |
|---|----------------------|--------------------|
| Patient demographics | | |
| <i>n</i> (females, males) | 177 (117,60) | 123 (94, 29) |
| Age (<i>M, SD</i>) | 37.0 (13.64) | 37.46 (12.34) |
| Therapist demographics | | |
| <i>n</i> (females, males) | 51 (35,16) | 21 (13,8) |
| Age (<i>M, SD</i>) | 50.73 (7.52) | 50.10 (8.33) |
| Years of professional experience (<i>M, SD</i>) | 16.92 (9.65) | 15.71 (10.16) |
| Certified as a supervisor? (%) | 39.4% | 35.7% |
| Consulting with a supervisor? (%), "regularly," "occasionally," "not at all" | 81.8% / 15.2% / 3.0% | 64.3% / 35.7% / 0% |
| Patient—therapist ratio (range) | 3.47 (1–10) | 5.86 (1–25) |
| Treatment characteristics | | |
| Number of sessions per therapy (<i>M, SD</i>) | 63 (45.43) | 53 (44.63) |
| Frequency of sessions (weekly, biweekly) | 88.1% / 9.0% | 69.3% / 28.4% |
| Number of sessions with guided imagery/ hypnotic trances per therapy (<i>M, SD</i>) | 10 (8.01) | 34 (34.55) |
| Number of therapies still ongoing at T6 (<i>n, %</i>) | 23 (13.0%) | 4 (3.3%) |
| End of treatment (completed, prematurely) | 59.4% / 40.6% | 79.3% / 20.6% |
| Diagnoses (<i>n, %</i>) | | |
| Somatic symptom disorder (PHQ) | 53 (29.9%) | 32 (26.0%) |
| Any mood disorder (PHQ) | 87 (49.2%) | 52 (42.3%) |
| Major depressive disorder | 55 (31.1%) | 38 (30.9%) |
| Other depressive disorder | 32 (18.1%) | 14 (11.4%) |
| Any anxiety disorder (PHQ) | 57 (32.2%) | 51 (41.4%) |
| Panic disorder | 34 (19.2%) | 21 (17.1%) |
| Other anxiety disorder | 23 (13.0%) | 30 (24.4%) |
| Any eating disorder (PHQ) | 19 (10.7%) | 13 (10.5%) |
| Binge eating disorder | 14 (7.9%) | 11 (8.9%) |
| Bulimia nervosa | 5 (2.8%) | 2 (1.6%) |
| Probable alcohol abuse (PHQ) | 34 (19.2%) | 15 (12.2%) |
| Cluster A personality disorder (ADP-IV) | 4 (2.3%) | 5 (4.1%) |
| Cluster B personality disorder (ADP-IV) | 24 (13.6%) | 11 (8.9%) |
| Cluster C personality disorder (ADP-IV) | 24 (13.6%) | 16 (13.0%) |
| Chronic somatic illness | 76 (44.4%) | 50 (41.7%) |
| Other treatment (<i>n, %</i>) | | |
| Current psychiatric medication | 102 (57.6%) | 75 (61.0%) |
| Previous psychological therapy | 108 (62.8%) | 70 (58.8%) |

Patient-Reported Diagnostic Assessment Measures

PRIME-MD Patient Health Questionnaire (PHQ). The PHQ (Spitzer, Kroenke, & Williams, 1999; German version: Löwe, Spitzer, Zipfel, & Herzog, 2002) is the fully self-administered version of the PRIME-MD. It is a screening instrument for a number of mental disorders according to the *DSM-IV* (axis I). The complete version as administered here contains 78 items of different scaling. Categorical diagnoses were computed using the *DSM-IV-TR* algorithm. The psychometric properties of the German version of the PHQ, including criterion validity when compared to

SCID-data, were found to be good (Gräfe, Zipfel, Herzog, & Löwe, 2004). In addition to the categorical diagnoses, the PHQ also contains three dimensional subscales. In this sample, internal consistencies for these were good: PHQ-9 Depression Severity, Cronbach's alpha = .86; GAD-7 Anxiety Scale, Cronbach's alpha = .84; PHQ-15 Somatic Symptom Severity, Cronbach's alpha = .81.

Assessment of *DSM-IV* Personality Disorders (ADP-IV). The ADP-IV (Schotte, De Doncker, Vankerckhoven, Vertommen, & Co-syns, 1998; German version: Doering et al., 2007) is a 94-item self-report instrument assess-

ing the *DSM-IV* personality disorders (axis II). Each item addresses “trait” as well as “distress/impairment” characteristics of a *DSM-IV* criterion. First, participants indicate on a 7-point scale (range: 1–7) how typical a certain personality trait is for them (sample item: “I absolutely cannot bear the idea that someone would leave or abandon me: therefore, I will do anything to prevent this.”). If a trait is judged as typical (≥ 5), participants then also indicate on a 3-point scale (range: 1–3) whether this characteristic has ever caused them or others distress or problems. For the present sample, each of the 10 *DSM-IV* personality disorders was judged as present or absent for each patient using the manualized cut-offs. Absolute and relative frequencies for the three DSM clusters of personality disorders were then computed. The German version of the ADP-IV has satisfactory correlations with SCID-diagnoses and expert ratings (Doering et al., 2007). It has also been shown to have very high sensitivity albeit relatively low specificity (Renn et al., 2008). In this sample, internal consistency for the 10 relevant subscales was $.66 \leq$ Cronbach’s alpha $\leq .84$.

Patient-Reported Clinical Outcome Measures

Brief Symptom Inventory (BSI). The BSI (Derogatis, 1993; German version: Franke, 2000) is the shortened version of the Symptom-Checklist-90-R. It is a 53-item self-report instrument where each item is rated on a five-point scale (range: 0–4) indicating distress or bother from psychopathological symptoms in the last seven days. The overall mean was computed to gain the GSI, an indicator of overall symptom distress. The BSI has been shown to have good construct validity. Test–retest reliability was $r_{xx} = .90$ (Franke, 2000). Internal consistency in the present sample was good to excellent ($.81 \leq$ Cronbach’s alpha $\leq .98$).

Inventory of Interpersonal Problems (IIP). The IIP (Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988; German version: Horowitz, Strauß, & Kordy, 2000) is a 64-item self-report measure. On a five-point scale (range: 0–4) participants indicate the degree to which they experience certain difficulties with significant others (sample item: “It is hard for me to take instructions from people who have authority over me”). The mean score was computed and interpreted as a global measure of interpersonal

problems. There is considerable evidence supporting the construct validity of the IIP (Gurtman, 1996; Horowitz et al., 2000) and it has been shown to be sensitive to change through psychotherapy (Huber, Henrich, & Klug, 2007). Average test–retest reliability was $r_{xx} = .86$. Internal consistency in the present sample was acceptable to excellent ($.77 \leq$ Cronbach’s alpha $\leq .96$).

Short Form of the Inventory of Personality Organization (IPO-16). The IPO-16 (Zimmermann et al., 2013) is a 16-item self-report instrument derived from the Inventory of Personality Organization (Lenzenweger, Clarkin, Kernberg, & Foelsch, 2001), which in turn is based on Kernberg’s model of personality organization. On a 5-point scale (range: 1–5) participants indicate the extent to which a given statement is true for them (sample item: “I find myself doing things which feel okay while I am doing them but which I later find hard to believe I did.”). The IPO’s original three-factor structure could successfully be replicated for the IPO-16 (Zimmermann et al., 2013). The mean score is commonly interpreted as a measure of general personality dysfunction. The IPO-16’s psychometric properties were found to be generally good (Zimmermann et al., 2013). Test–retest reliability was $r_{xx} = .85$. In the present sample internal consistency was good (Cronbach’s alpha_{T1} = .84, Cronbach’s alpha_{T6} = .85).

Questions on Life Satisfaction (FLZ^M). The FLZ^M (Henrich & Herschbach, 2000) is a 32-item self-report instrument. It is organized into two modules (each with eight items addressing different aspects): “General Life Satisfaction” (General-LS; sample items: “income/financial security,” “family life/children”) and “Satisfaction with Health” (Health; sample items: “physical condition/fitness,” “independence from help/care”). Participants rate each item twice: once for subjective importance of that aspect and once for their current satisfaction regarding that aspect. Both ratings are then combined into a “weighted satisfaction” score. The sums of these scores for both subscales provide global measures for general and health-related life satisfaction respectively (range: –96–160). The FLZ^M has been shown to yield good psychometric properties, including sensitivity to change (Henrich & Herschbach, 2000). Test-retest reliabilities were $r_{xx} = .87$ for Gen-

eral-LS and $r_{xx} = .85$ for Health (Henrich & Herschbach, 2000). In the present sample internal consistency ranged from questionable to good for General-LS ($.66 \leq$ Cronbach's $\alpha \leq .84$) and from acceptable to good for Health ($.79 \leq$ Cronbach's $\alpha \leq .86$).

Psychological Mindedness Scale (PMS). The PMS (Conte, Ratto, & Karasu, 1996) is a 45-item self-report instrument assessing psychological mindedness. In the PMS, participants indicate on a 4-point scale (range: 1–4) the degree to which a given statement is true for them (sample item: “Talking about your worries to another person helps you to understand your problems better”). Conte et al., (1996) reported good psychometric properties. Test–retest reliability was $r_{xx} = .92$. The sum of all (partially reverse-keyed) item-scores is usually interpreted as a general measure of psychological mindedness (Shill & Lumley, 2002; range: 45–180). In the present sample, internal consistency was good for T1 (Cronbach's $\alpha_{T1} = .86$) as well as for T6 (Cronbach's $\alpha_{T6} = .86$).

Sociodemographic information and health economic outcome measures. In addition to sociodemographic information, the questionnaire also included items on how frequently the patients had seen a doctor or underwent other medical procedures such as surgeries or endoscopies during the last year, how many days of sick leave they had in the last year, whether they are on medication, whether they had recent hospital stays, and how much inpatient or outpatient psychotherapy they had had prior to the current treatment.

Therapist-Reported Clinical Outcome Measures

Clinical Global Impression—Severity (CGI-S). The CGI-S (Guy, 1976) is one of two items making up the Clinical Global Impression. The clinician is asked to indicate the severity of the patient's psychopathology on a 7-point scale (range: 1–7) while taking into account all available information about the patient as well as the sum of the practitioner's clinical experience. The second item, asking for the amount of a patient's improvement, was not used in the present study because it could not be shown to have any advantages over a prepost measurement using only the severity item (Forkmann et al., 2011). As for validity, the CGI-S has been shown to correlate moderately

to strongly with longer disorder-specific rating scales for depression and anxiety disorders (e.g., Bandelow, Baldwin, Dolberg, Andersen, & Stein, 2006) as well as psychotic disorders. Interrater reliability was reported to be good: $ICC_{1,1} = .89$ (Honer et al., 1995, as cited in Johnson, 2010).

Global Assessment of Functioning (GAF). The GAF is listed as axis V in the *DSM-IV-TR* (APA, 2000, German version: Saß, Wittchen, Zaudig, & Houben, 2003). The clinician is asked to indicate their impression of the overall social, occupational, and psychological functioning of a patient. The GAF consists of one numeric scale (range: 1–100) accompanied by a detailed guideline on rating patient functioning. The GAF has repeatedly been criticized for problems with both predictive and concurrent validity (Aas, 2010). It was nevertheless included here because of its brevity and ease of use. Also, interrater reliability was found to range from satisfactory to good ($ICC_{1,1} = .84$; Honer et al., 1995, as cited in Johnson, 2010)—at least when raters were experienced clinicians (Söderberg, Tungström, & Armelius, 2005).

Adherence Measures

Comparative Psychotherapy Process Scale—Therapist Form (CPPS). The CPPS (Hilsenroth, Blagys, Ackerman, Bonge, & Blais, 2005) is an instrument designed to assess techniques and interventions in psychotherapies. It can be completed by therapists, patients, or external raters and contains 20 items to be rated on a 7-point scale: 10 items describing therapist techniques and interventions typical for cognitive-behavioral therapies and 10 items typical for psychodynamic therapies. The CPPS has been shown to have good reliability and to discriminate well between psychodynamic and nonpsychodynamic forms of therapy (Hilsenroth et al., 2005). For the present study, only the therapist version was used after the end of treatment. The therapists were asked to indicate how characteristic each item was for their work with this particular patient, rather than for their therapeutic approach in general.⁶

⁶ One of our reviewers pointed out that it would have improved the value of the CPPS data if we had asked the therapists explicitly to review all of their clinical notes before filling out the questionnaire. We recommend this for future studies.

Internal consistency in the present sample was acceptable: Cronbach's $\alpha_{CB} = .78$, Cronbach's $\alpha_{PI} = .71$. As part of the validation of the CPPS, Hilsenroth et al. (2005) had experienced psychodynamic and CBT clinicians give prototype ratings of ideally conducted sessions in their theoretical orientation. These prototype ratings were used as a baseline with which to compare how strongly the therapists in the present study adhered to their respective methods and how typical the interventions and techniques were for different theoretical orientations.

Process Scale for Imagery- and Trance-Based Psychodynamic Therapies—Therapist Form (PITT). As assimilative integrations, GIP and HY are forms of psychodynamic therapy but also employ imagery- and trance-based techniques not associated with the traditional psychodynamic approach as measured in the CPPS. To assess these techniques as well, we included an additional instrument for imagery- and trance-based methods, the Prozessskala für imaginations- und trancebasierte tiefenpsychologische Methoden (PITT; Sell, Schöpfer-Mader, Brömmel, & Möller, in press). The PITT uses the same rating scale and item formats as the CPPS and was developed to be used in conjunction with the latter. It contains 6 items specific to GIP (sample item: “I instructed my patient to imagine a certain motif—e.g. a meadow, a mountain, etc.—while in a state of relaxation”) and 7 items specific to HY (sample item: “I encouraged my patient to imagine different parts of their personality, name them and bring them into dialogue with each other.”). As with the CPPS, only the therapist version was used after the end of treatment. Internal consistency in the present sample was good for the HY subscale (Cronbach's $\alpha = .86$) but questionable for the GIP subscale (Cronbach's $\alpha = .65$). Since prototype ratings for GIP and HY had not yet been available, we established a baseline by asking the study therapists to indicate at the beginning of each treatment which techniques they plan on using with each of their patients.

Treatment Procedures

Both GIP and HY are considered to be forms of integrative psychodynamic therapy. Both methods use imagery- and trance-based techniques respectively within their psychodynamic

framework. The treatment strategy for GIP, following an initial diagnostic phase, was described by Dieter (2001) as follows: While in a state of mild relaxation the patient is instructed to day-dream—often about a particular motif. Everything that comes up within the patient's imagination shall be directly reported to the therapist. During the imagination, patient and therapist are then meant to engage in a dialogue in order to further explore the patient's inner world. According to Kottje-Birnbacher (1992), a typical treatment in GIP will involve several sessions with such guided imagery (experiencing), but there will also be sessions in which the content of the prior imaginations is meant to be understood and worked through therapeutically (processing). This working through is meant to follow basic principles of psychodynamic therapy and involves working with transference, countertransference, defense, and resistance (Pichler, 2011).

According to Kanitschar (2009), a typical HY treatment involves an initial phase of diagnostics and working on the patient's ability to enter into hypnotic trances. The latter are conceived of as relaxed waking-states that are characterized among other things by a heightened suggestibility and a narrowed focus of attention (Harrer, 2008). For the subsequent working phase, three modes of hypnotherapeutic intervention can be distinguished: “building and actualizing of a patient's resources,” “working through conflicts and offering corrective emotional experiences,” and “future-orientation with a focus on possible solutions.” As in GIP, this working phase is meant to rely on the principles of psychodynamic therapy.

Treatment characteristics are summarized in Table 1. The difference between GIP and HY in mean numbers of session was marginally significantly, $t(210) = 1.89$, $p = .06$, $d = .22$. Compared to HY, GIP treatments had a higher percentage of weekly rather than biweekly or less frequent sessions, $t(210) = 4.03$, $p < .001$. The percentage of sessions involving guided imagery or hypnotic trances respectively was significantly lower for GIP than for HY, $t(210) = 8.58$, $p > .001$. This difference is in line with both methods' theoretical foundations. Whereas Bahrke and Nohr (2013) state that guided imagery is customarily used in one out of four sessions, Kanitschar (H. Kanitschar, personal communication, October 10, 2016) says

that hypnotic trances are used in more than half of the sessions of a typical HY treatment. The percentage of therapies that were still ongoing after the study duration of 30 months was higher for GIP than for HY, $t(210) = 2.88, p = .004$. The percentage of therapies ending prematurely was significantly higher for GIP than for HY treatments, $t(210) = 3.01, p = .003$. This could be because on average GIP treatments were slightly longer and thereby more expensive for the patients. The most common reasons for terminating therapy early were “no longer being able to pay for therapy,” “dissatisfaction with therapy,” and “having moved to a different area.”

Since both methods are not manualized, the CPPS and PITT were used as indicators of adherence. Table 2 shows mean scores for the techniques and interventions employed in GIP and HY therapies as rated by the therapists at the end of each treatment. Multiple comparisons (Šidák corrected) showed that—like the ideal psychodynamic session—both GIP, $t(78) = 12.66, p < .001, d = 2.88$, and HY therapists, $t(59) = 3.26, p = .001, d = .86$, reported on average to have used more psychodynamic interventions than CBT interventions. Also—as in the respective planned treatments—GIP therapists used more GIP interventions than HY interventions, $t(78) = 8.71, p < .001, d = 1.95$, and vice versa, $t(59) = 2.39, p = .01, d = .62$. It is noteworthy that HY treatments were more eclectic than GIP treatments: The amount of CBT technique used by HY therapists was higher than that used by GIP therapists, $t(139) = 9.82, p < .001, d = 1.66$, but lower than in the ideal CBT sessions, $t(71) = 1.81, p = .04, d = .73$. Furthermore, there was significantly more GIP technique used in HY treatments than HY technique used in GIP treatments, $t(139) = 10.90, p < .001, d = 1.89$. Overall, these self-report data suggest that GIP as well as HY treatments were conducted in adherence to the tenets of the psychodynamic tradition in general as well as to the two methods' specific treatment technique.

Statistical Analysis

All analyses were performed with IBM SPSS Statistics for Windows, Version 22.0. We used an alpha level of .05 for all statistical tests. Marginally significant results ($p < .10$) are also

reported. Following an exploratory data analysis, we imputed single missing items using the EM algorithm. For every item the percentage of missing values was $< 3\%$. For each of the six measurement points, Little's MCAR test confirmed that the data were missing completely at random. We then used a series of hierarchical linear models (HLMs) to assess change from T1 to T6 for all outcome measures as well as the progression of change over all six measurement points for symptom distress. The estimation method was Maximum Likelihood, and we chose an autoregressive covariance structure, AR(1). The nesting of client data within therapists was accounted for in all analyses: Measurement occasions were at level 1, individual patients at level 2, and therapists at level 3. According to the recommendations by Tasca and Gallop (Gallop & Tasca, 2009; Tasca & Gallop, 2009), we developed the models sequentially, building from a base model (random-intercept-only model) and adding growth terms and covariates (“number of sessions” and “therapy completed at T6”) in subsequent steps. We centered all predictors to the grand mean. We calculated $\sim R^2$ scores for each full model relative to the respective base model. Following Feingold (2009), we calculated the effect size d from the model-estimated marginal means using the standard deviation at T1. We then corrected the effect size for the pretest–posttest correlation using the formula provided by Morris and DeShon (2002). All analyses were based on the completer sample. Since BSI data were available for all six measurement points, we also performed an intention-to-treat analysis for the BSI data using the last observation carried forward (LOCF) method. To test whether there was any relationship between therapeutic technique and treatment outcome, we computed Pearson correlations (Šidák corrected) between CPPS/PITT scores and regression residuals of the HLMs.

We followed the procedure suggested by Jacobson and Truax (1991) to test for clinically significant change. We computed the Reliable Change Index (RCI; Jacobson & Truax, 1991) for each participant on each outcome scale. We then defined each participant's status as either “recovered” (T6 score fell into the functional range and corresponded to reliable change), “improved” (T6 score fell into the dysfunctional range but corresponded to reliable change), “un-

Table 2
Employed Interventions and Therapeutic Techniques [Range: 0–6]

| Type of therapy | Employed techniques and interventions | | | |
|---|---|--|--|--|
| | Psychodynamic technique (CPPS) <i>M (SD)</i> | Cognitive Behavioral technique (CPPS) <i>M (SD)</i> | Guided Imagery technique (PITT) <i>M (SD)</i> | Hypnotherapy technique (PITT) <i>M (SD)</i> |
| Guided Imagery treatments (GIP, this sample, <i>n</i> = 80) | 4.39 (.70) | 1.98 (.98) | 4.63 (2.01) | .97 (1.74) |
| Hypnotherapy treatments (HY, this sample, <i>n</i> = 61) | 5.03 (.93) | 4.00 (1.46) | 3.93 (1.39) | 4.71 (1.13) |
| Ideal Cognitive Behavioral session (Hilsenroth et al., 2005, <i>n</i> = 13) | 3.05 (.61) | 4.75 (.61) | — | — |
| Ideal psychodynamic session (Hilsenroth et al., 2005, <i>n</i> = 30) | 5.04 (.57) | 1.43 (.61) | — | — |
| Planned GIP treatments (this sample, <i>n</i> = 177) | 4.86 (.62) | 2.14 (.89) | 5.29 (.75) | 1.47 (1.48) |
| Planned HY treatments (this sample, <i>n</i> = 123) | 4.95 (.67) | 4.01 (1.17) | 4.55 (1.00) | 5.19 (.71) |

Note. CPPS = Comparative Psychotherapy Process Scale; PITT = Process Scale for Imagery and Trance-Based Psychodynamic Therapies.

changed” (no reliable change), or “deteriorated” (reliable change in the negative direction).

We defined dysfunctional and functional ranges using the following samples from the literature: For the BSI, we used a nonclinical sample by Franke (2000) and a sample of German psychotherapy outpatients (Geisheim et al., 2002) to define a cut-off point at $c = .73$. For the IIP, we used a representative sample of the German population (Brähler, Horowitz, Kordy, Schumacher, & Strass, 1999) and a combination of four clinical samples (Brähler et al., 1999) to define a cut-off point at $c = 1.53$. For the IPO-16, we used a representative norm sample of the general German population (Zimmermann, Benecke, Hörz-Sagstetter, & Dammann, 2015) and a German psychotherapy outpatient sample (Zimmermann et al., 2013) to define a cut-off point at $c = 2.00$. For FLZ^M, we used a representative sample of the German population (Henrich & Herschbach, 2000) and a sample of psychosomatic-psychotherapy inpatients assessed in the late 1990s in Germany (G. Henrich, personal communication, February 6, 2016; $N_{\text{General-LS}} = 1044$, $M_{\text{General-LS}} = 9.49$, $SD = 33.66$, $N_{\text{Health}} = 1027$, $M_{\text{Health}} = -6.60$, $SD = 35.76$). The cut-off points were defined at $c = 34.99$ for General-LS and at $c = 33.90$ for Health. For the PMS, the CGI-S, and the GAF no adequate norm samples

were available. We therefore used $+2 SD$ from the mean at T1 as a criterion for clinically significant change. For the health economic items there were neither norm samples nor reliability measures available. Here we used the method of percent improvement ($\geq 50\%$ improvement from baseline and $\geq 25\%$ individual change on the entire range of the scale) as an indicator of clinically significant change (Hiller, Schindler, & Lambert, 2012).

Results

Change in Clinical Measures After 30 Months

Descriptive statistics for all clinical outcome measures are presented in Table 3 and Table 4. For GIP treatments, HLM analyses revealed statistically significant change after 30 months for every outcome measure considered while controlling for “number of sessions” and “therapy completed until T6.” Table 3 summarizes the results. Participants improved with regard to symptom distress (BSI GSI, $d = .78$), interpersonal problems (IIP total score, $d = .39$), personality organization (IPO-16 total score, $d = .50$), general life satisfaction (FLZM-General-

Table 3
Change in Clinical Measures After 30 Months for Guided Imagery Psychotherapy (GIP): HLMs (ML, Controlled for “Number of sessions” and “Therapy Completed Until T6”)

| Outcome variable | Change after 30 months | | | | | Reliable change and clinically significant change | | | | |
|---|--|---------------------------------------|-------|-------|------------|---|------------|------------|------------|------------|
| | T1: Pre treatment ^a M (SD) | T6: +30 Months ^b M (SD) | t | p | $\sim R^2$ | d | ++ (n, %) | + | o (n, %) | - (n, %) |
| Patient-reported clinical outcome measures | | | | | | | | | | |
| BSI GSI [Range: 0–4] | 1.05 (.72) | .51 (.49) | 5.52 | <.001 | .52 | .78 | 19 (31.7%) | 3 (5.0%) | 38 (63.3%) | 0 (0%) |
| BSI GSI (LOCF) | 1.05 (.72) | .76 (.65) | 2.40 | .009 | .47 | .47 | — | — | — | — |
| IIP total [Range: 0–4] | 1.37 (.58) | 1.15 (.53) | 2.84 | .003 | .36 | .39 | 13 (21.3%) | 2 (3.3%) | 44 (72.1%) | 2 (3.3%) |
| IPO-16 total [Range: 1–5] | 2.11 (.58) | 1.74 (.44) | 3.74 | <.001 | .50 | .50 | 12 (19.7%) | 3 (4.9%) | 46 (75.4%) | 0 (0%) |
| FLZ ^M , General-LS [Range: –96–160] | 24.81 (34.90) | 43.50 (31.37) | 3.16 | .002 | .32 | .64 | 12 (19.7%) | 2 (3.3%) | 44 (72.1%) | 3 (4.9%) |
| FLZ ^M , Health [Range: –96–160] | 31.58 (36.40) | 63.03 (37.18) | 6.74 | <.001 | .47 | .63 | 19 (31.1%) | 1 (1.6%) | 41 (67.2%) | 0 (0%) |
| PMS total [Range: 4 – 180] | 138.35 (14.17) | 142.22 (14.59) | 2.19 | .02 | .09 | .71 | 1 (1.6%) | 10 (16.4%) | 44 (72.1%) | 6 (9.8%) |
| Health economic outcome measures | | | | | | | | | | |
| Doctors' visits last year | 11.13 (8.87) | 6.41 (7.00) | 3.53 | <.001 | .14 | .48 | 27 (49.1%) | 9 (16.4%) | 6 (10.9%) | 13 (23.6%) |
| Days of sick leave last year | 21.25 (49.07) | 7.61 (16.37) | 2.00 | .03 | .19 | .27 | 21 (35.6%) | 1 (1.7%) | 21 (35.6%) | 16 (27.1%) |
| Days in hospital last year | 3.07 (11.66) | .77 (1.91) | 2.52 | .01 | .13 | .12 | 12 (19.7%) | 0 (0%) | 41 (67.2%) | 8 (13.1%) |
| Therapist-reported clinical outcome measures | | | | | | | | | | |
| CGI-S [Range: 1–7] | 4.54 (.82) | 3.04 (1.55) | 11.61 | <.001 | .51 | 1.76 | 46 (37.7%) | 45 (36.9%) | 25 (20.5%) | 6 (4.9%) |
| GAF [Range: 1–100] | 54.28 (11.48) | 71.34 (14.93) | 13.81 | <.001 | .63 | 1.46 | 46 (36.8%) | 28 (22.4%) | 48 (38.4%) | 3 (2.4%) |

Note. $\sim R^2$ -scores are for the full model relative to the respective base model (random-intercept-only model); *d*-scores are the differences between estimated marginal means, standardized by the SD at T1 corrected for the pretest–posttest correlation (Morris & DeShon, 2002); ++ = “recovered”; + = “improved”; o = “unchanged”; - = “deteriorated.”
^a *n* = 177; ^b *n* = 61; ^c *n* = 177; ^d *n* = 131.

Table 4
Change in Clinical Measures After 30 Months for Hypnotherapy (HY): HLMs (ML, Controlled for “Number of Sessions” and “Therapy Completed Until T6”)

| Outcome variable | Change after 30 months | | | | Reliable change and clinically significant change | | | | | |
|---|---------------------------------------|------------------------------------|----------|----------|---|----------|------------|------------|------------|-----------|
| | T1: Pre treatment ^a M (SD) | T6: +30 Months ^b M (SD) | <i>t</i> | <i>p</i> | $\sim R^2$ | <i>d</i> | ++ (n, %) | +(n, %) | o (n, %) | – (n, %) |
| Patient-reported clinical outcome measures | | | | | | | | | | |
| BSI GSI [Range: 0–4] | 1.03 (.74) | .52 (.51) | 4.43 | <.001 | .34 | .88 | 12 (33.3%) | 7 (19.4%) | 17 (47.2%) | 0 (0%) |
| BSI GSI (LOCF) | 1.03 (.74) | .75 (.68) | 2.52 | .007 | .64 | .50 | — | — | — | — |
| IIP total [Range: 0–4] | 1.35 (.59) | 1.00 (.59) | 3.51 | <.001 | .22 | .67 | 17 (47.2%) | 0 (0%) | 16 (44.4%) | 3 (8.3%) |
| IPO-16 total [Range: 1–5] | 1.99 (.52) | 1.72 (.53) | 5.62 | <.001 | .52 | .94 | 8 (22.2%) | 1 (2.8%) | 27 (75.0%) | 0 (0%) |
| FLZ ^M -General-LS [Range: –96–160] | 29.97 (35.35) | 59.39 (46.07) | 4.72 | <.001 | .33 | .83 | 11 (30.6%) | 0 (0%) | 24 (66.7%) | 1 (2.8%) |
| FLZ ^M -Health [Range: –96–160] | 31.74 (44.91) | 62.10 (53.83) | 4.31 | <.001 | .38 | .81 | 9 (25.0%) | 1 (2.8%) | 26 (72.2%) | 0 (0%) |
| PMS total [Range: 45–180] | 139.08 (13.10) | 140.47 (13.62) | 2.18 | .02 | .06 | .46 | 0 (0%) | 12 (33.3%) | 20 (55.6%) | 4 (11.1%) |
| Health economic outcome measures | | | | | | | | | | |
| Doctors' visits last year | 13.50 (15.28) | 7.75 (10.91) | 3.38 | .001 | .14 | .47 | 14 (45.2%) | 5 (16.1%) | 7 (22.6%) | 5 (16.1%) |
| Days of sick leave last year | 19.99 (40.20) | 8.42 (15.94) | 1.77 | .045 | .18 | .26 | 15 (12.2%) | 3 (10.0%) | 5 (16.7%) | 7 (23.3%) |
| Days in hospital last year | 5.39 (16.07) | .78 (1.51) | 1.37 | .09 | .11 | .48 | 8 (23.5%) | 0 (0%) | 20 (58.8%) | 6 (17.6%) |
| Therapist-reported clinical outcome measures | | | | | | | | | | |
| CGI-S [Range: 1–7] | 4.43 (1.06) | 2.64 (1.46) | 11.23 | <.001 | .57 | 1.58 | 36 (45.0%) | 30 (37.5%) | 10 (12.5%) | 4 (5.0%) |
| GAF [Range: 1–100] | 58.98 (12.37) | 76.19 (13.28) | 11.27 | <.001 | .62 | 1.38 | 28 (35.0%) | 19 (23.8%) | 33 (41.3%) | 0 (0%) |

Note. $\sim R^2$ -scores are for the full model relative to the respective base model (random-intercept-only model); *d*-scores are the differences between estimated marginal means, standardized by the SD at T1 corrected for the pretest–posttest correlation (Morris & DeShon, 2002); ++ = “recovered”; + = “improved”; o = “unchanged”; – = “deteriorated.”
^a *n* = 123; ^b *n* = 36; ^c *n* = 123; ^d *n* = 81.

LS, $d = .64$), health-related life satisfaction (Health, $d = .63$), psychological mindedness (PMS total score, $d = .71$), doctors' visits during the last year ($d = .48$), days of sick leave during the last year ($d = .27$), days in hospital during the last year ($d = .12$), therapist-rated symptom severity (CGI-S, $d = 1.76$), and therapist-rated global functioning (GAF, $d = 1.46$). The intention-to-treat analysis (LOCF) for the BSI data also showed significant improvement but, as expected, with a smaller effect size ($d = .47$) than in the completer sample.

Regarding clinically significant change in GIP treatments (see Table 3 for details) on the patient-reported clinical outcome measures, between 18% and 36.7% of the participants experienced positive change (either "recovered" or "improved"), between 63.3% and 75.4% of the participants remained unchanged, and between 0% and 9.8% deteriorated. On the health economic outcome measures, between 19.7% and 65.5% of the participants experienced positive change, between 10.9% and 67.2% of the participants remained unchanged, and between 13.1% and 27.1% deteriorated. On the therapist-reported clinical outcome measures, between 59.2% and 74.6% of the participants had positive change, between 20.5% and 38.4% of the participants remained unchanged, and between 2.4% and 4.9% deteriorated.

For HY treatments, HLM analyses showed statistically significant change for all but one of the outcome measures when controlling for "number of sessions" and "therapy completed until T6." The results are summarized in Table 4. Participants improved with regard to symptom distress (BSI GSI, $d = .88$), interpersonal problems (IIP total score, $d = .67$), personality organization (IPO-16 total score, $d = .95$), general life satisfaction (FLZM-General, $d = .83$), health-related life satisfaction (FLZM-Health, $d = .81$), psychological mindedness (PMS total score, $d = .46$), doctors' visits during the last year ($d = .47$), days of sick leave during the last year ($d = .26$), therapist-rated symptom severity (CGI-S, $d = 1.58$), and therapist-rated global functioning (GAF, $d = 1.38$). The decrease of number of days in hospital during the last year did not reach significance in the HLM, despite the effect being of medium size ($d = .48$). The intention-to-treat analysis (LOCF) for the BSI

data showed significant improvement as well but again with a smaller effect size ($d = .50$).

Regarding clinically significant change in HY treatments (see Table 4 for details) on the patient-reported clinical outcome measures, between 25% and 52.7% of the participants experienced positive change, between 44.4% and 75.0% of the participants remained unchanged, and between 0% and 11.1% deteriorated. On the health economic outcome measures, between 22.2% and 61.3% of the participants experienced positive change, between 16.7% and 58.8% of the participants remained unchanged, and between 16.1% and 23.3% deteriorated. On the therapist-reported clinical outcome measures, between 58.8% and 82.5% of the participants had positive change, between 12.5% and 41.3% of the participants remained unchanged, and between 0% and 5% deteriorated.

Relationship Between Therapeutic Technique and Treatment Outcome

For GIP, the only statistically significant correlations between therapeutic technique (CPPS; PITT) and change in outcome variables were for the therapist-reported clinical outcome measures: More prominent use in psychodynamic technique was associated with greater change in both CGI-S, $r = .36$, $p = .005$, and GAF, $r = .27$, $p = .04$. However, only the relationship between CGI-S and psychodynamic technique remained significant after Šidák correction. For HY, no significant correlations between CPPS/PITT and change in outcome variables were found. Overall, the data thus reveals little evidence for certain specific types of interventions or techniques explaining a significant portion of the variance in therapeutic outcome.

Progression of Change in Symptom Distress

To examine at what points in time during the 30 months after the beginning of treatment the change in symptom distress (BSI GSI) occurred, we ran HLMs over six points of measurement (T1-T6; one every 6 months), again controlling for "number of sessions" and "therapy completed until T6." The results are displayed in Table 5. For the completer sample of GIP treatments, a linear term, $t(273) = 3.46$, $p < .001$, and a quadratic term,

Table 5
Change in Symptom Distress Over Time for Guided Imagery Psychotherapy (GIP) and Hypnopsychotherapy (HY): HLMs (ML, Controlled for "Number of Sessions" and "Therapy Completed Until T6")

| Outcome variable | T1 Pre treatment M (SD) | T2 + 6 months M (SD) | T3 + 12 Months M (SD) | T4 + 18 Months M (SD) | T5 + 24 Months M (SD) | T6 + 30 Months M (SD) | $\sim R^2$ | Šidák Post-hoc analysis (d) |
|----------------------|----------------------------|-------------------------|-----------------------------|--------------------------|-----------------------------|-----------------------------|------------|--------------------------------------|
| GIP | | | | | | | | |
| BSI GSI [Range: 0–4] | 1.05 (.72) ^a | .71 (.54) ^b | .69 (.64) ^c | .58 (.56) ^d | .60 (.60) ^e | .51 (.49) ^f | .32 | T1 > T2 (d = .52); T3 > T6 (d = .56) |
| BSI GSI (LOCF) | 1.05 (.72) | .83 (.64) | .82 (.68) | .80 (.68) | .80 (.69) | .76 (.65) | .15 | T1 > T2 (d = .36) |
| HY | | | | | | | | |
| BSI GSI [Range: 0–4] | 1.03 (.74) ^g | .76 (.64) ^h | .70 (.66) ⁱ | .66 (.64) ^j | .57 (.59) ^k | .52 (.51) ^l | .52 | T1 > T2 (d = .51); T2 > T6 (d = .55) |
| BSI GSI (LOCF) | 1.03 (.74) | .85 (.69) | .78 (.69) | .79 (.70) | .76 (.68) | .75 (.68) | .03 | T1 > T2 (d = .36) |

Note. $\sim R^2$ -scores are for the full model relative to the respective base model (random-intercept-only model).

^a n = 177; ^b n = 115; ^c n = 94; ^d n = 82; ^e n = 82; ^f n = 73; ^g n = 123; ^h n = 68; ⁱ n = 56; ^j n = 48; ^k n = 41; ^l n = 36.

$t(265) = 1.88, p < .05$, were statistically significant, indicating that the decrease in symptom distress follows a quadratic growth curve. Šidák post hoc analyses for the model-estimated marginal means revealed mean levels of symptom distress to have significantly decreased from T1 to T2, $t(304) = 3.32, p < .001, d = .52$, and then again from T3 to T6, $t(284) = 2.47, p < .01, d = .56$. No significant differences were found in between. In the intention-to-treat analysis (LOCF), the linear term was only marginally significant $t(302) = 1.54, p = .06$, and the quadratic term was significant, $t(384) = 1.91, p = .03$. Šidák post hoc analyses for the model-estimated marginal means showed that for the LOCF-sample symptom distress had significantly decreased from T1 to T2, $t(595) = 3.01, p < .001, d = .36$. The difference between T2 and T6 was only marginally significant, $t(595) = 1.43, p = .08, d = .15$, and no significant differences were found in between.

For HY, analogous results were found in the completer sample: The linear effect, $t(123) = 4.75, p < .001$, as well as the quadratic polynomial, $t(132) = 3.56, p < .001$, were significant. Šidák Post hoc analyses showed the estimated marginal means to have decreased between T1 and T2, $t(112) = 2.87, p < .01, d = .51$, and in this case between T2 and T6, $t(95) = 1.95, p < .05, d = .55$. Again, no significant differences were found in between. The intention-to-treat analysis (LOCF) also showed the linear term, $t(182) = 2.01, p = .02$, and the quadratic term, $t(264) = 2.92, p = .002$, to be significant. As for GIP, Šidák post hoc analyses for the model-estimated marginal means showed that for the LOCF-sample symptom distress had only significantly decreased from T1 to T2, $t(380) = 2.41, p < .001, d = .36$. The difference between T2 and T6 was not significant, $t(380) = 1.37, p = .13, d = .18$, and neither were any differences in between.

For both methods, these results indicate an initial rapid gain within the first 6 months after the beginning of treatment and also a trend toward a second slower gain ongoing until at least 30 months after the beginning of treatment. Figure 2 illustrates this progression by plotting the descriptive means in symptom distress (BSI GSI; completer sample) over T1 to T6.

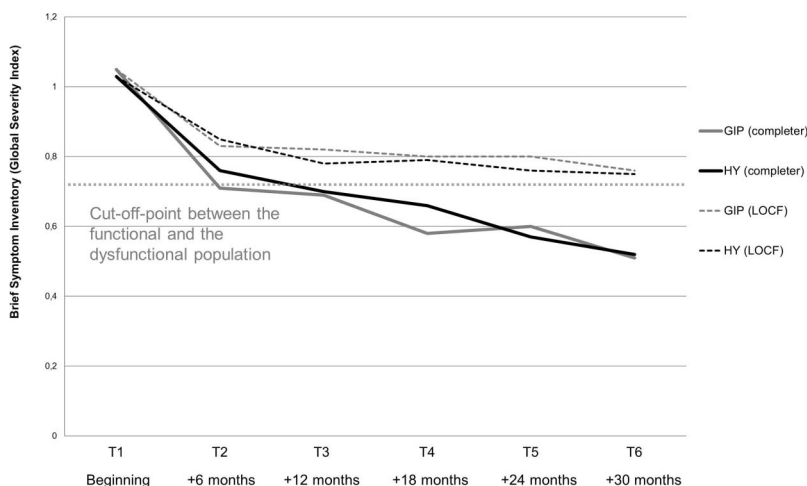


Figure 2. Descriptive means in symptom distress over time in Guided Imagery Psychotherapy (GIP) and hypnotherapy (HY).

Discussion

This study aimed to assess the effectiveness of two outpatient methods—Guided-Imagery Psychotherapy (GIP) and Hypnotherapy (HY)—both integrative imagery- and trance-based psychodynamic therapies. An ecologically valid sample was recruited, reflecting treatment modalities and patient characteristics as they actually occur within a public health care system—in this case Austria. Compared to previous studies on these methods, sample size and study duration were comparable or better. A diverse selection of outcome measures was implemented, comprising symptom-related and more general scales, relationship- and personality-oriented methods, as well as health economic measures. Regarding adherence, we found evidence that therapists in both methods conducted their treatments in accordance with basic psychodynamic principles as well as the tenets of the specific trance- and imagery-based components integrated within their core framework.

As an overall result, 30 months after the beginning of treatment participants in both methods showed significant improvement on all outcome measures except days in hospital for HY. Both immediate well-being, as assessed through symptom distress and life satisfaction, as well as the quality of the participants' relationships and their level of personality organi-

zation increased over the course of treatment. Psychological mindedness, which had previously been shown to be associated with greater therapeutic benefit from these therapies, increased after treatment as well. Furthermore, three (two out of three for HY) health economic measures indicated that participants who had undergone treatment used fewer public health services and had fewer days of sick leave than prior to their treatment. From the beginning to the end of treatment, therapist-rated global functioning and severity improved as well.

In terms of the extent of change, the mean within-group effect sizes for therapist-rated outcome (GIP, $d = 1.61$; HY, $d = 1.48$) were within the range of the meta-analysis by Leichsenring and Leibing (2003, $\bar{d} = 1.79$). The effects we found for change in patient-reported symptom distress (GIP, $d = .78$; HY, $d = .88$) also were comparable to change in symptom distress as reported in meta-analyses on psychodynamic short-term therapies ($\bar{d} = .90$, Leichsenring, Rabung, & Leibing, 2004), psychodynamic long-term therapies ($\bar{d} = .89$, Leichsenring & Rabung, 2008, 2011), as well as psychodynamic therapies for personality disorders ($\bar{d} = .81$, Leichsenring & Leibing, 2003). The meta-analysis on clinical hypnosis by Revenstorf (2006) also found effects of a similar overall effect size ($\bar{d} = .83$). With regard to change in personality organization (GIP, $d =$

.50; HY, $d = .94$), the average effects in GIP and HY are smaller than for measures of personality and psychosocial functioning in the meta-analysis on psychoanalytic treatments by de Maat et al. (2013, $\bar{d} = 1.08$). The studies included in their meta-analysis, however, had a session frequency of twice-weekly or higher as compared to weekly or biweekly in GIP and HY. The effects we found on interpersonal problems (GIP, $d = .39$; HY, $d = .67$) were smaller than those found for GIP by Sachsse et al., 2016, $d = .81$), and those we found for life satisfaction (GIP, $\bar{d} = .64$, HY, $\bar{d} = .82$) were smaller than those found for GIP by von Wietersheim et al. (2003, $d = .97$).

We need to add, however, that the two latter studies examined samples of German patients whose therapy was fully covered by statutory health insurances. In comparison, the patients in the present sample often had to pay for part or all of their treatment themselves (see footnote 2). Because of this socioeconomically restricted access to therapy, severely ill patients are likely to be less common within the Austrian sample. When comparing preintervention scores, patients in the two German samples have significantly higher distress ratings than the Austrian patients in the current sample on all comparable measures. The postintervention scores on the other hand were just as good or better in the Austrian sample. In other words, the average participant in the present sample began treatment in a more subclinical range with comparatively low levels of distress and therefore had less room for positive change.

Examining the rates of change for GIP and HY therapies revealed that relatively few patients accounted for the effects described above. Between 44.4% and 75.4% of patients did not have reliable change. This seems to be because many of these patients had not been severely ill to begin with, rather than having significant pathology that did not respond to treatment: According to the patient-reported clinical outcome measures, 42.7% of the patients were already in the range of the functional population when commencing treatment and between 75.4% and 80.6% of the patients were in the functional range 30 months after the beginning of treatment. We should add, however, that more than half of the patient sample had been in therapy before, and almost half of the patients suffer from at least one chronic somatic illness. This could be an indicator that with regard to

psychological problems, the sample may not have been highly impaired but chronic. To an extent this would also explain fairly long treatments and slower progress.

Comparably for both methods, repeated assessment of symptom distress over six points of measurement showed that improvement with regard to this scale can—at least for the complete sample—be described in two consecutive steps: A rapid initial gain within the first 6 months after the beginning of treatment and a second, slower improvement over the remaining 18 to 24 months. This can be seen as evidence of a prolonged process of change during these therapies. The rate of change decreases over time but the participants nonetheless continue to improve—at least until 30 months after the beginning of treatment if not longer, ultimately leading the participants into the range of the functional population.

Several factors need to be considered that limit the study's explanatory power. First, the lack of a control group diminishes internal validity. A sample of participants who have problems comparable to the treatment sample yet are not treated by a psychotherapist but only by a GP would seem most appropriate for a truly naturalistic study. Considerable efforts were made to recruit such a sample—unfortunately, without sufficient success. Furthermore, internal validity is compromised because we were only able to obtain two points of measurement for most of the outcome measures. While the number of sessions and whether the treatment had been completed was controlled for statistically, there is still no way to discriminate between end-of-treatment effects and follow-up effects. In addition, the investigated treatments are fairly heterogeneous regarding diagnoses. While this is in line with the naturalistic approach, it does not allow for any disorder-specific results—especially with the sample being too small to be stratified. Also, as mentioned earlier, access to psychotherapy in Austria is to varying degrees determined by place of residence and socioeconomic status. Therefore, a limitation concerning external validity is that the sample is likely to be representative of people seeking psychotherapy in Austria but most likely not representative of all people in need of psychotherapeutic help within that community.

A key limitation is the considerable number of dropouts. On the one hand, dropout rates are

within the range of comparable naturalistic outpatient studies (e.g., Ward, Wood, & Awal, 2013) and dropout analyses revealed no significant differences between the intention-to-treat sample and completers on the preintervention measures. On the other hand, however, we still cannot rule out the possibility of dropouts systematically biasing the sample in an undeterminable way and thereby skewing the results. The dropouts also led to a fairly small N in the completer sample. The low statistical power caused by the small N could in turn be an explanation for the surprisingly small number of significant correlations between therapeutic technique and outcome.

Making sense of the dropouts is difficult because of the relatively few measurement points. However, the data do allow for some hypotheses as to who dropped out and why (Behar & Borkovec, 2003). First, regarding the high level of attrition between T1 and T2, the patients received the initial questionnaire directly from their therapist, whereas thereafter all assessment was done by postal service. We presume that the former was associated with a higher level of commitment for many patients. Furthermore, the data show that dropout rates were about twice as high (39%) for patients who had already completed treatment than for those still in therapy. The fact that patients were less likely to remain in the study after their therapy had ended suggests that the final results underestimate the overall effectiveness of the treatments. On the other hand, we also found evidence that dissatisfaction with treatment was related to study dropouts: Patients who indicated to have ended therapy prematurely because of “dissatisfaction with therapy” (22%) had a three times higher chance of dropping out (66%) than the remainder of the sample. This suggests that treatments that were not going well from a patient perspective had a lower likelihood of being included in the final results. Most likely this leads to an overestimation of treatment effectiveness. To conclude, the high number of dropouts has the potential to distort the results in different directions. Since we have no means of quantifying these distortions more precisely, further research of the effectiveness of GIP and HY is needed in order for us to gain more confidence in the exact sizes of the treatments’ effects.

In spite of the limitations discussed above, the results of the study indicate that treatments in both

GIP and HY are effective under the naturalistic conditions of the Austrian health care system. From the perspective of psychotherapy integration, this also suggests that investigating smaller integrative traditions within the context of public health care delivery settings is promising. While newly developed integrative therapies are often very well thought through and based on recent research, their implementation into routine practice often takes a long time or does not happen at all. By also looking for integrative components in established approaches that are already in place as part of national health care delivery systems, we gain the opportunity to study creative integrative attempts—such as the use of imagery interventions within a psychodynamic framework—in routine practice, with different types of patients, treatment lengths, and therapeutic environments.

References

- Aas, I. H. M. (2010). Global Assessment of Functioning (GAF): Properties and frontier of current knowledge. *Annals of General Psychiatry, 9*, 20. <http://dx.doi.org/10.1186/1744-859X-9-20>
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, 4th edition, text revision (DSM-IV-TR)*. Washington, DC: Author.
- Arkowitz, H. (1997). Integrative theories of therapy. In P. L. Wachtel & S. B. Messer (Eds.), *Theories of psychotherapy: Origins and evolution* (pp. 227–288). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10239-006>
- Arntz, A. (2011). Imagery rescripting for personality disorders. *Cognitive and Behavioral Practice, 18*, 466–481. <http://dx.doi.org/10.1016/j.cbpra.2011.04.006>
- Arntz, A. (2012). Imagery rescripting as a therapeutic technique: Review of clinical trials, basic studies, and research agenda for personality disorders. *Journal of Experimental Psychopathology, 3*, 189–208. <http://dx.doi.org/10.5127/jep.024211>
- Bahrke, U., & Nohr, K. (2013). *Katathym Imaginative Psychotherapie: Lehrbuch der Arbeit mit Imaginationen in psychodynamischen Psychotherapien* [Guided imagery psychotherapy: Textbook on working with imagination in psychodynamic psychotherapies]. Berlin, Germany: Springer. <http://dx.doi.org/10.1007/978-3-642-03254-7>
- Bandelow, B., Baldwin, D. S., Dolberg, O. T., Andersen, H. F., & Stein, D. J. (2006). What is the threshold for symptomatic response and remission for major depressive disorder, panic disorder, social anxiety disorder, and generalized

- anxiety disorder? *The Journal of Clinical Psychiatry*, 67, 1428–1434. <http://dx.doi.org/10.4088/JCP.v67n0914>
- Barber, J. P., Muran, C., McCarthy, K. S., & Keefe, J. R. (2013). Research on dynamic therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 443–494). Hoboken, NJ: Wiley.
- Behar, E. S., & Borkovec, T. D. (2003). Psychotherapy outcome research. In I. B. Irving (Series Ed.), J. A. Schinker, W. F. Velive (Vol. Eds.), *Handbook of psychology: Vol. 2, Research methods in psychology* (pp. 214–240). Hoboken, NJ: Wiley. <http://dx.doi.org/10.1002/0471264385.wei0209>
- Brähler, E., Horowitz, L. M., Kordy, H., Schumacher, J., & Strauss, B. (1999). Zur Validierung des Inventars zur Erfassung Interpersonaler Probleme (IIP). Ergebnisse einer Repräsentativbefragung in Ost- und Westdeutschland [Validation of the inventory for Interpersonal Problems (IIP): Results of a representative study]. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 49, 422–431.
- Brewin, C. R., Wheatley, J., Patel, T., Fearon, P., Hackmann, A., Wells, A., . . . Myers, S. (2009). Imagery rescripting as a brief stand-alone treatment for depressed patients with intrusive memories. *Behaviour Research and Therapy*, 47, 569–576. <http://dx.doi.org/10.1016/j.brat.2009.03.008>
- Brown, D., & Fromm, E. (1986). *Hypnotherapy and hypnoanalysis*. Mahwah, NJ: Erlbaum.
- Castonguay, L. G., Newman, M. G., Borkovec, T. D., Holtforth, M. G., & Maramba, G. G. (2005). Cognitive-behavioral assimilative integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 241–260). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/med:psych/9780195165791.003.0011>
- Conte, H. R., Ratto, R., & Karasu, T. B. (1996). The Psychological Mindfulness Scale: Factor structure and relationship to outcome of psychotherapy. *The Journal of Psychotherapy Practice and Research*, 5, 250–259.
- DeFife, J., Drill, R., Beinashowitz, J., Ballantyne, L., Plant, D., Smith-Hansen, L., . . . Nakash, O. (2015). Practice-based psychotherapy research in a public health setting: Obstacles and opportunities. *Journal of Psychotherapy Integration*, 25, 299–312. <http://dx.doi.org/10.1037/a0039564>
- de Maat, S., de Jonghe, F., de Kraker, R., Leichsenring, F., Abbass, A., Luyten, P., . . . Dekker, J. (2013). The current state of the empirical evidence for psychoanalysis: A meta-analytic approach. *Harvard Review of Psychiatry*, 21, 107–137.
- Derogatis, L. R. (1993). *BSI Brief Symptom Inventory: Administration, scoring, and procedure manual*. Minneapolis, MN: National Computer Systems.
- Dieter, W. (2001). Die Katathym Imaginative Psychotherapie—Eine tiefenpsychologische Behandlungsmethode [Guided Imagery Psychotherapy: A depth psychological method]. *Imagination*, 23, 5–41.
- Doering, S., Renn, D., Höfer, S., Rumpold, G., Smrekar, U., Janecke, N., . . . Schüssler, G. (2007). Validierung der deutschen Version des Fragebogens zur Erfassung von DSM-IV Persönlichkeitsstörungen (ADP-IV) [Validation of the German version of the Assessment of DSM-IV Personality Disorders (ADP-IV) Questionnaire]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 53, 111–128. <http://dx.doi.org/10.13109/zptm.2007.53.2.111>
- Edwards, D. (2007). Restructuring implicational meaning through memory-based imagery: Some historical notes. *Journal of Behavior Therapy and Experimental Psychiatry*, 38, 306–316. <http://dx.doi.org/10.1016/j.jbtep.2007.10.001>
- Erickson, M. H., & Rossi, E. L. (1981). *Experiencing hypnosis: Therapeutic approaches to altered states*. New York, NY: Irvington.
- Esplen, M. J., Garfinkel, P. E., Olmsted, M., Gallop, R. M., & Kennedy, S. (1998). A randomized controlled trial of guided imagery in bulimia nervosa. *Psychological Medicine*, 28, 1347–1357. <http://dx.doi.org/10.1017/S0033291798007405>
- Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, 14, 43–53. <http://dx.doi.org/10.1037/a0014699>
- Forkmann, T., Scherer, A., Boecker, M., Pawelzik, M., Jostes, R., & Gauggel, S. (2011). The Clinical Global Impression Scale and the influence of patient or staff perspective on outcome. *BMC Psychiatry*, 11, 83. <http://dx.doi.org/10.1186/1471-244X-11-83>
- Forrester, M. (2000). *Psychology of the image*. London, England: Routledge. <http://dx.doi.org/10.4324/9780203446928>
- Franke, G. H. (2000). *BSI: Brief Symptom Inventory—German Version*. Göttingen, Germany: Beltz.
- Fromm, E. (1977). An ego-psychological theory of altered states of consciousness. *International Journal of Clinical and Experimental Hypnosis*, 25, 372–387. <http://dx.doi.org/10.1080/00207147708415992>
- Gallop, R., & Tasca, G. A. (2009). Multilevel modeling of longitudinal data for psychotherapy researchers: II. The complexities. *Psychotherapy Research*, 19, 438–452. <http://dx.doi.org/10.1080/10503300902849475>
- Geisheim, C., Hahlweg, K., Fiegenbaum, W., Frank, M., Schröder, B., & von Witzleben, I. (2002). Das

- Brief Symptom Inventory (BSI) als Instrument zur Qualitätssicherung in der Psychotherapie [The German version of the Brief Symptom Inventory (BSI): Reliability and validity in a sample of outpatient psychotherapy patients]. *Diagnostica*, 48, 28–36. <http://dx.doi.org/10.1026/0012-1924.48.1.28>
- Gilbert, P. (2010). *Compassion focused therapy*. New York, NY: Routledge.
- Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem Gesundheitsfragebogen für Patienten (PHQ-D) [Screening for psychiatric disorders with the Patient Health Questionnaire (PHQ): Results from the German validation study]. *Diagnostica*, 50, 171–181. <http://dx.doi.org/10.1026/0012-1924.50.4.171>
- Gurtman, M. B. (1996). Interpersonal problems and the psychotherapy context: The construct validity of the Inventory of Interpersonal Problems. *Psychological Assessment*, 8, 241–255. <http://dx.doi.org/10.1037/1040-3590.8.3.241>
- Guy, W. (1976). *ECDEU assessment manual for psychopharmacology*. Rockville, MD: US Department of Health, Education, and Welfare Public Health Service Alcohol, Drug Abuse, and Mental Health Administration.
- Hackmann, A. (2005). Compassionate imagery in the treatment of early memories in Axis I anxiety disorders. In P. Gilbert (Ed.), *Compassion: Conceptualisations, research and use in psychotherapy* (pp. 352–368). Hove, England: Routledge.
- Hagl, M. (2013). Zur Wirksamkeit von Hypnose und Hypnotherapie. Eine Studienübersicht für die Jahre 2010–2012 [Efficacy of clinical hypnosis—Intervention studies from 2010 to 2012]. *Hypnose-ZHH*, 8, 145–181.
- Hagl, M. (2014). Studien zur Wirksamkeit von Klinischer Hypnose und Hypnotherapie im Jahr 2013 [Efficacy research in the field of clinical hypnosis in the year 2013]. *Hypnose-ZHH*, 9, 147–168.
- Hall, E., Hall, C., Stradling, P., & Young, D. (2006). *Guided imagery: Creative interventions in counselling & psychotherapy*. London, England: Sage.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose–response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9, 329–343. <http://dx.doi.org/10.1093/clipsy.9.3.329>
- Harrer, M. (2008). Wirkkonzepte in der Hypnosepsychotherapie [Why is hypnopsychotherapy effective?]. *Imagination*, 30, 5–37.
- Henrich, G., & Herschbach, P. (2000). Questions on Life Satisfaction (FLZ^M)—A short questionnaire for assessing subjective quality of life. *European Journal of Psychological Assessment*, 16, 150–159. <http://dx.doi.org/10.1027//1015-5759.16.3.150>
- Hiller, W., Schindler, A. C., & Lambert, M. J. (2012). Defining response and remission in psychotherapy research: A comparison of the RCI and the method of percent improvement. *Psychotherapy Research*, 22, 1–11. <http://dx.doi.org/10.1080/10503307.2011.616237>
- Hilsenroth, M. J., Blagys, M. D., Ackerman, S. J., Bonge, D. R., & Blais, M. A. (2005). Measuring psychodynamic-interpersonal and cognitive-behavioral techniques: Development of the comparative psychotherapy process scale. *Psychotherapy: Theory, Research, Practice, Training*, 42, 340–356. <http://dx.doi.org/10.1037/0033-3204.42.3.340>
- Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review*, 30, 349–362. <http://dx.doi.org/10.1016/j.cpr.2010.01.001>
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56, 885–892. <http://dx.doi.org/10.1037/0022-006X.56.6.885>
- Horowitz, L. M., Strauß, B., & Kordy, H. (2000). *Inventar zur Erfassung interpersonaler Probleme—Deutsche Version* [Inventory of Interpersonal Problems (IIP): German Version]. Göttingen, Germany: Beltz.
- Huber, D., Henrich, G., & Klug, G. (2007). The Inventory of Interpersonal Problems (IIP): Sensitivity to change. *Psychotherapy Research*, 17, 474–481. <http://dx.doi.org/10.1080/10503300600856160>
- 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. <http://dx.doi.org/10.1037/0022-006X.59.1.12>
- Johnson, D. L. (2010). *A compendium of psychosocial measures: Assessment of people with serious mental illness in the community*. New York, NY: Springer.
- Kanitschar, H. (2009). Hypnosepsychotherapie, ein integratives, tiefenpsychologisch fundiertes Verfahren [Hypnopsychotherapy, an integrative psychodynamic approach]. *Hypnose-ZHH*, 4, 1–23.
- Kottje-Birnbacher, L. (1992). Strukturierende Faktoren des Katathymen Bilderlebens [Structuring factors of guided imagery]. *Praxis der Psychotherapie und Psychosomatik*, 37, 164–173.
- Leichsenring, F., & Leibling, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: A meta-analysis. *The American Journal of Psychiatry*, 160, 1223–1232. <http://dx.doi.org/10.1176/appi.ajp.160.7.1223>
- Leichsenring, F., & Rabung, S. (2008). Effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *Journal of the American Medical*

- Association*, 300, 1551–1565. <http://dx.doi.org/10.1001/jama.300.13.1551>
- Leichsenring, F., & Rabung, S. (2011). Long-term psychodynamic psychotherapy in complex mental disorders: Update of a meta-analysis. *The British Journal of Psychiatry*, 199, 15–22. <http://dx.doi.org/10.1192/bjp.bp.110.082776>
- Leichsenring, F., Rabung, S., & Leibing, E. (2004). The efficacy of short-term psychodynamic psychotherapy in specific psychiatric disorders: A meta-analysis. *Archives of General Psychiatry*, 61, 1208–1216. <http://dx.doi.org/10.1001/archpsyc.61.12.1208>
- Lenzenweger, M. F., Clarkin, J. F., Kernberg, O. F., & Foelsch, P. A. (2001). The Inventory of Personality Organization: Psychometric properties, factorial composition, and criterion relations with affect, aggressive dyscontrol, psychosis proneness, and self-domains in a nonclinical sample. *Psychological Assessment*, 13, 577–591. <http://dx.doi.org/10.1037/1040-3590.13.4.577>
- Leuner, H. (1984). *Guided affective imagery: Mental imagery in short-term psychotherapy*. New York, NY: Thieme-Stratton.
- Levy, K. N., Ehrenthal, J. C., Yeomans, F. E., & Caligor, E. (2014). The efficacy of psychotherapy: Focus on psychodynamic psychotherapy as an example. *Psychodynamic Psychiatry*, 42, 377–421. <http://dx.doi.org/10.1521/pdps.2014.42.3.377>
- Löwe, B., Spitzer, R. L., Zipfel, S., & Herzog, W. (2002). *PHQ-D: Gesundheitsfragebogen für Patienten. Manual: Komplettversion und Kurzform* [PHQ-D: Patient Health Questionnaire: German manual]. Karlsruhe, Germany: Pfizer.
- Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105–125. <http://dx.doi.org/10.1037/1082-989X.7.1.105>
- Norcross, J. C. (2005). A primer on psychotherapy integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 3–23). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/med:psych/9780195165791.003.0001>
- O’Leary, E., & Barry, N. (2006). Gestalt reminiscence therapy. In E. O’Leary & M. Murphy (Eds.), *New approaches to integration in psychotherapy* (pp. 50–60). New York, NY: Routledge.
- Orlinksy, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy: Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior Change* (4th ed., pp. 270–376). New York, NY: Wiley.
- Pichler, M. (2011). Katathym Imaginative Psychotherapie [Guided Imagery Psychotherapy]. In G. Stumm (Ed.), *Psychotherapie, Schulen und Methoden: Eine Orientierungshilfe für Theorie und Praxis* [Psychotherapy, orientations and methods: A guide for theory and practice] (pp. 109–117). Vienna, Austria: Falter Verlag.
- Renn, D., Höfer, S., Schüssler, G., Rumpold, G., Smrekar, U., Janecke, N., & Doering, S. (2008). Dimensionale Diagnostik mit dem Fragebogen zur Erfassung von DSM-IV-Persönlichkeitsstörungen (ADP-IV) [A dimensional diagnostic approach using the Assessment of DSM-IV Personality Disorders (ADP-IV) questionnaire]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 54, 212–226. <http://dx.doi.org/10.13109/zptm.2008.54.3.212>
- Revenstorf, D. (2006). Expertise zur Beurteilung der wissenschaftlichen Evidenz des Psychotherapieverfahrens Hypnotherapie [Review of the empirical evidence of hypnotherapy]. *Hypnose-ZHH*, 1, 7–164.
- Robertson, D. (2012). *The practice of cognitive-behavioural hypnotherapy: A manual for evidence-based clinical hypnosis*. London, England: Karnac.
- Roeckelein, J. E. (2004). *Imagery in psychology: A reference guide*. London, England: Praeger.
- Sachsse, U., Imruck, B. H., & Bahrke, U. (2016). Evaluation ambulanter Behandlungen mit Katathym Imaginativer Psychotherapie KIP: Eine naturalistische Studie [Evaluation of outpatient psychotherapy with Guided Imagery Psychotherapy: A naturalistic outcome study]. *Ärztliche Psychotherapie*, 11, 87–92.
- Saß, H., Wittchen, H.-U., Zaudig, M., & Houben, I. (2003). *Diagnostisches und statistisches Manual psychischer Störungen, DSM-IV-TR, Textrevision DSM-IV*, German version]. Göttingen, Germany: Hogrefe.
- Schmucker, M. R., & Dancu, C. F. (1999). *Cognitive-behavioral treatment for adult survivors of childhood trauma: Imagery, rescripting and reprocessing*. New York, NY: Rowman & Littlefield.
- Schotte, C. K. W., de Doncker, D., Vankerckhoven, C., Vertommen, H., & Cosyns, P. (1998). Self-report assessment of the DSM-IV personality disorders. Measurement of trait and distress characteristics: The ADP-IV. *Psychological Medicine*, 28, 1179–1188. <http://dx.doi.org/10.1017/S0033291798007041>
- Sell, C., Möller, H., & Taubner, S. (2016). Symptomreduktion und Prädiktoren des Behandlungserfolgs bei Katathym Imaginativer Psychotherapie und Hypnosepsychotherapie: Ergebnisse nach einem Jahr unter Praxisbedingungen [Symptom reduction and predictors of treatment success in Guided Imagery Psychotherapy and Hypnotherapy: Results after one year under naturalistic conditions]. Manuscript submitted for publication.

- Sell, C., Schöpfer-Mader, E., Brömmel, B., & Möller, H. (in press). Behandlungstechnik in Katathym Imaginativer Psychotherapie, Hypnose-psychotherapie und Autogener Psychotherapie: Die Therapeutenversion der 'Prozessskala für imaginations- und trancebasierte tiefenpsychologische Methoden' (PITT) [Treatment technique in Guided Imagery Psychotherapy, Hypnotherapy, and Autogenic Psychotherapy: The therapist version of the 'Process Scale for Imagery- and Trance-Based Psychodynamic Therapies' (PITT)]. *Psychotherapie Forum*.
- Shedler, J. (2012). The efficacy of psychodynamic psychotherapy. In R. A. Levy, J. S. Ablon, & H. Kächele (Eds.), *Psychodynamic Psychotherapy Research: Evidence-Based Practice and Practice-Based Evidence* (pp. 9–25). New York, NY: Humana Press. http://dx.doi.org/10.1007/978-1-60761-792-1_2
- Shill, M. A., & Lumley, M. A. (2002). The Psychological Mindfulness Scale: Factor structure, convergent validity, and gender in a non-psychiatric sample. *Psychology and Psychotherapy: Theory, Research and Practice*, 75, 131–150. <http://dx.doi.org/10.1348/147608302169607>
- Smit, Y., Huibers, M. J., Ioannidis, J. P., van Dyck, R., van Tilburg, W., & Arntz, A. (2012). The effectiveness of long-term psychoanalytic psychotherapy—A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, 32, 81–92. <http://dx.doi.org/10.1016/j.cpr.2011.11.003>
- Söderberg, P., Tungström, S., & Armelius, B. A. (2005). Reliability of global assessment of functioning ratings made by clinical psychiatric staff. *Psychiatric Services*, 56, 434–438. <http://dx.doi.org/10.1176/appi.ps.56.4.434>
- Spitzer, R. L., Kroenke, K., & Williams, J. B. W. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Journal of the American Medical Association*, 282, 1737–1744. <http://dx.doi.org/10.1001/jama.282.18.1737>
- Stigler, M., & Pokorny, D. (2001). Emotions and primary process in guided imagery psychotherapy: Computerized text-analytic measures. *Psychotherapy Research*, 11, 415–431. <http://dx.doi.org/10.1093/ptr/11.4.415>
- Stigler, M., & Pokorny, D. (2012). Eine Dekade der KIP-Prozessforschung im Überblick [A decade of process-research on GIP: An overview]. In H. Ullmann & E. Wilke (Eds.), *Handbuch Katathym Imaginative Psychotherapie* [Handbook of guided imagery psychotherapy] (pp. 122–145). Bern, Switzerland: Huber.
- Stricker, G. (2010). A second look at psychotherapy integration. *Journal of Psychotherapy Integration*, 20, 397–405. <http://dx.doi.org/10.1037/a0022037>
- Stricker, G., & Gold, J. (2005). Assimilative psychodynamic psychotherapy. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (pp. 221–240). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/med:psych/9780195165791.003.0010>
- Tasca, G. A., & Gallop, R. (2009). Multilevel modeling of longitudinal data for psychotherapy researchers: I. The basics. *Psychotherapy Research*, 19, 429–437. <http://dx.doi.org/10.1080/10503300802641444>
- von Wietersheim, J., Wilke, E., Röser, M., & Meder, G. (2003). Ergebnisse der Katathym-imaginativen Psychotherapie: Die Effektivität der Katathym-imaginativen Psychotherapie in einer ambulanten Längsschnittstudie [The effects of Guided Affective Imagery in a longitudinal outpatient study]. *Psychotherapeut*, 48, 173–178.
- Ward, A., Wood, B., & Awal, M. A. (2013). Naturalistic psychodynamic psychotherapy study: Evaluating outcome with a patient perspective. *British Journal of Psychiatry*, 29, 292–314.
- Watkins, J. G. (1992). *Hypnoanalytic techniques. The practice of clinical hypnosis* (Vol. 2). New York, NY: Irvington.
- West, W. (2000). Eclecticism and integration in humanistic therapy. In S. Palmer & R. Woolfe (Eds.), *Integrative and eclectic counselling and psychotherapy* (pp. 218–232). London, England: Sage. <http://dx.doi.org/10.4135/9781446280409.n12>
- Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. New York, NY: Guilford Press.
- Zimmermann, J., Benecke, C., Hörz, S., Rentrop, M., Peham, D., Bock, A., . . . Dammann, G. (2013). Validierung einer deutschsprachigen 16-Item-Version des Inventars der Persönlichkeitsorganisation (IPO-16) [Validity of a German 16-item version of the Inventory of Personality Organization (IPO-16)]. *Diagnostica*, 59, 3–16. <http://dx.doi.org/10.1026/0012-1924/a000076>
- Zimmermann, J., Benecke, C., Hörz-Sagstetter, S., & Dammann, G. (2015). Normierung der deutschsprachigen 16-Item-Version des Inventars der Persönlichkeitsorganisation (IPO-16) [Standardization of a German 16-item-version of the Inventory of Personality Organization (IPO-16)]. *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 61, 5–18. <http://dx.doi.org/10.13109/zptm.2015.61.1.5>
- Zinbarg, R. E., Craske, M. G., & Barlow, D. H. (2006). *Mastery of your anxiety and worry: Therapist guide* (2nd ed.). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/med:psych/9780195300024.001.0001>

Received August 31, 2016

Revision received January 2, 2017

Accepted January 11, 2017 ■