The Emerging Evidence for Long-Term Psychodynamic Therapy

Falk Leichsenring, Allan Abbass, Patrick Luyten, Mark Hilsenroth, and Sven Rabung

Abstract: There is growing evidence from RCTs supporting the efficacy of both short-term (STPP) and long-term psychodynamic psychotherapy (LTPP) for specific mental disorders. In a first series of meta-analyses, LTPP was shown to be superior to shorter forms of psychotherapy, especially in complex mental disorders. However, the evidence for LTPP has not gone unchallenged. After several responses have addressed the raised concerns, a recent meta-analysis by Smit and colleagues (2012) again challenges the efficacy of LTPP.

Method: From a methodological perspective, a critical analysis of the Smit et al. meta-analysis was performed. Furthermore, we conducted two new meta-analyses adding studies not included in previous meta-analyses. The purpose was to examine whether the results of the previous meta-analyses are stable.

Results: Due to differing inclusion criteria, the meta-analysis by Smit et al. actually compared LTPP to other forms of *long-term* psychotherapy. Thus, they essentially showed that LTPP was as efficacious as other forms of *long-term* therapy. For this reason the meta-analysis by Smit et al. does not question the results of previous meta-analyses showing that LTPP is superior to shorter forms of psychotherapy. In addition, the Smit et al. meta-analysis was shown to suffer from several methodological shortcomings. The new meta-analyses we performed did not find significant deviations from previous results. In complex mental disorders LTPP proved to be significantly superior to shorter forms of therapy corroborating results of previous meta-analyses.

Conclusions: Data on dose-effect relations suggest that for many patients with complex mental disorders, including chronic mental disorders and personality disorders, short-term psychotherapy is not sufficient. For these patients, long-

---

Falk Leichsenring, D.Sc., Professor, Clinic of Psychosomatics and Psychotherapy, University of Giessen, Germany. Allan Abbass, H.D., Professor, Dalhousie University Halifax, Nova Scotia, Canada. Patrick Luyten, Ph.D., Professor, Department of Psychology, University of Leuven, Belgium, and Research Department of Clinical, Educational, and Health Psychology, University College London. Mark J. Hilsenroth, Ph.D., Professor, The Derner Institute of Advanced Psychological Studies, Adelphi University, Garden City, New York.

Sven Rabung, Ph.D., Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, and Department of Psychology, Alpen-Adria-Universität Klagenfurt, Austria.

Psychodynamic Psychiatry, 41(3) 361–384, 2013
© 2013 The American Academy of Psychoanalysis and Dynamic Psychiatry

term treatments may be indicated. The meta-analyses presented here provide further support for LTPP in these populations. Nevertheless, there is a need for more research in LTPP and other long-term psychotherapies.

There is a growing body of evidence from randomized controlled trials (RCTs) for the efficacy of psychodynamic therapy in specific mental disorders (Abbass, Hancock, Hernderson, & Kisley, 2006; Gerber et al., 2011; Leichsenring, Kruse, & Rabung, in press). This is true for both short-term psychodynamic psychotherapy (STPP) and long-term psychodynamic psychotherapy (LTPP; Abbass, Hancock, Hernderson, & Kisley, 2004, 2006; Gerber et al., 2011; Leichsenring, 2009; Leichsenring et al., in press; Leichsenring & Rabung, 2008, 2011b). Further, the observed benefits of psychodynamic psychotherapy increase in long-term follow-up suggesting the treatment brings structural changes enabling delayed and continued gains (Town, Diener, Abbass, Leichsenring, Driessen, & Rabung, 2012).

Several recent meta-analyses specifically addressing the efficacy and effectiveness of LTPP (de Maat, Philipszoon, Schoevers, Dekker, & de Jonghe, 2007; de Maat, de Jonghe, Schoevers, & Dekker, 2009; Leichsenring & Rabung, 2008, 2011b) suggest that LTPP is superior to shorter forms of psychotherapy in complex mental disorders, that is, for chronic mental disorders, patients with multiple mental disorders or personality disorders (Leichsenring & Rabung, 2011b). This is consistent with data on dose-effect relations which suggest that for many patients with chronic mental disorders or personality disorders, short-term psychotherapy is not sufficient (Kopta, Howard, Lowry, & Beutler, 1994). For these patients, long-term treatments may be more effective. For LTPP a recent meta-analysis provided evidence in support of this assumption (Leichsenring & Rabung, 2011b), and this may be applicable to other forms of long-term psychotherapy as well. For patients with severe personality disorders, for example, long-term CBT is expected to be more efficacious than short-term CBT (Beck & Rush, 1995, p. 1854). Dialectical-behavior therapy (DBT) and schema-focused therapy (SFT), for instance, were developed as long-term treatments and proved to be efficacious in patients with borderline personality disorder (Leichsenring, Leibing, Kruse, New, & Leweke, 2011). Likewise, the response to brief psychotherapy is limited in many depressed patients (Cuijpers, van Straten, Bohlmeijer, Hollon, & Andersson, 2010; Cuijpers, van Straten, Schuurmans, et al., 2010) which has led to the development of treatment packages that offer maintenance and continuation treatment, particularly in patients with chronic depression (Steven, 2011).

Hence, there is growing support for the evaluation of, training in, and dissemination of long-term treatments, as many patients, and particularly those with chronic and complex psychological problems, may need more extended treatments.

Yet, evidence for the efficacy of LTPP—and of STPP as well—has not gone unchallenged (Bhar & Beck, 2009; Bhar et al., 2010). Several responses have addressed these concerns (Leichsenring & Rabung, 2009, 2011a, 2011b, 2012; Leichsenring et al., 2011). Moreover, these responses have raised concerns about potential investigator allegiance effects and the use of double standards in judging the quality of evidence for LTPP. For instance, trials of psychodynamic psychotherapy have been criticized for failing to achieve certain methodological standards (e.g., sufficient statistical power, ensurance of treatment integrity), whereas similar methodological shortcomings are typically ignored in interpreting studies of other treatment modalities (Leichsenring & Rabung, 2011a; Leichsenring et al., 2011). A recent quality-based review corroborated this view, showing that quality of RCTs of CBT for depression was similar to the quality of RCTs of psychodynamic psychotherapy (Thoma et al., 2011).

Smit and colleagues (2012), however, recently published a meta-analysis of LTPP which concluded (p. 81) that the evidence for LTPP is "limited and at best conflicting," and that the "recovery rate of various mental disorders was equal after LTPP or various control treatments, including treatment as usual." Hence, the conclusions of this meta-analysis head-on collide with those of previous meta-analyses of LTPP (Leichsenring & Rabung, 2008, 2011b).

In the following, we address Smit et al.'s conclusions in detail.[1] In addition, we present two new meta-analyses of LTPP which include studies that had not been included in previous meta-analyses, but were included by Smit et al. (2012). The purpose of these meta-analyses was to examine whether the results of previous meta-analyses are stable. Before addressing these issues, however, we feel that it is imperative to acknowledge the limitations of research concerning the effects of LTPP, including the relative paucity of well-conducted studies, the lack of standardization and manualization of treatments, and the need to improve the quality of RCTs more generally. However, these problems are not unique for LTPP. Notwithstanding these clear limitations, we feel that existing research deserves a more balanced discussion than provided by Smit et al. (2012). In the following we will discuss the me-

1. We regret that *Clinical Psychology Review*, the journal in which the Smit et al. meta-analyses was presented, did not provide the opportunity to comment on the meta-analysis by Smit et al. (2012).

ta-analysis of Smit et al. from a methodological perspective. After that, results of new meta-analyses will be presented.

I. INCLUSION CRITERIA AFFECT RESULTS

The results of any meta-analysis and the conclusions that can be drawn from it, heavily depend on the studies that are included and excluded. If it was the intention by Smit et al. (2012) to test whether the results of previous meta-analyses are valid, this necessarily implies the use of the same inclusion criteria. In the following we will show that Smit et al. used inclusion and exclusion criteria different from those used in prior meta-analyses of LTPP. For this reason, their results cannot be compared to those of previous meta-analyses.

(1) Definition of LTPP

There is no generally accepted "standard" duration for long-term psychodynamic psychotherapy. Consistent with experts in the field (Crits-Christoph & Barber, 2000), previous meta-analyses defined LTPP by at least 50 sessions of psychodynamic psychotherapy *or* by a psychodynamic treatment that lasts for one year or longer (Leichsenring & Rabung, 2008, 2011b). Smit et al. (2012), however, used the following definition (p. 82): "We defined long-term psychotherapy as having at least 40 sessions and continuing for at least one year." Thus, Smit et al. required two criteria to be met simultaneously, that is, at least 40 sessions *and* a treatment lasting for at least one year. Regarding their specification of the minimum number of sessions, Smit et al. referred to a Cochrane review for psychodynamic psychotherapy which defined short-term psychodynamic psychotherapy as less than 40 sessions (Abbass et al., 2004). However, this definition does not imply that a treatment of 40 or more sessions is long-term. It may be also regarded as medium-term.[2]

Applying the definition used by Smit et al., however, has the surprising consequence that two studies included by Smit et al. do not fulfill their own inclusion criteria. The study by Dare, Eisler, Russel, Treasure, and Dodge (2001), as the psychodynamic treatment in this study, consisted of only 24.9 sessions over a year of treatment. Similarly, the study by McMain et al. (2009) was included, although the average number

2. In sports, a 100 m race is a short-distance race. However, not every distance longer than 100 m is a long distance race, it may be a medium-distance race, for example, of 200 m or 400 m.

of sessions in the condition labelled by Smit et al. as "LTPP" was only 31 sessions (McMain et al., 2009, p. 1371). Thus, both of these studies, following their own definition of LTPP, should not have been included by Smit et al.

(2) Completed Treatments

Previous meta-analyses required that the treatments among included studies be completed and terminated (Leichsenring & Rabung, 2008, 2011b). It is questionable whether data from ongoing treatments provide valid estimates for treatment outcome at termination or follow-up, for example, if patients received only half of the "dose" of treatment when outcome is assessed.[3]

For this reason, several studies were excluded from previous meta-analyses of LTPP (Doering et al., 2010; Giesen-Bloo et al., 2006; Puschner, Kraft, Kächele, & Kordy, 2007). In the study by Giesen-Bloo et al., for example, 19 of 42 patients treated with LTPP (45%) were still in treatment when outcome was assessed, and only 2 patients had completed LTPP; 21 of 44 were lost to therapy and assessments (Giesen-Bloo et al., 2006). In the comparison group, 27 out of 44 patients (61%) were still in treatment, and only 6 patients had completed the treatment; 11 of 44 were lost to therapy and assessments. Smit et al. (2012), however, apparently did not require treatments to be terminated as they included the study by Giesen-Bloo. In this light, it is not clear why Smit et al. did not also include the RCT by Doering et al., showing that LTPP was superior to treatments conducted by experienced community therapists (Doering et al., 2010).

(3) Clear Representatives of LTPP Proper

In previous meta-analyses, only studies of LTPP consistent with definitions by experts in the field were included. Leichsenring and Rabung (2011b), for example, applied a definition by Gabbard (2004). Smit et al. (2012) apparently did not require such a definition and included stud-

3. By analogy, if one runner starts for a 100 m race and another one for a 10,000 m race, the time taken after 100 m will not be representative for the short-distance speed of the second runner. The runners will adapt their speed to the short versus long distance they are going to face. This is true for patients in psychotherapy as well (Knekt et al., 2008). Psychotherapy is not a drug that works equally under different conditions, but a psychosocial process.

ies that are clearly not representative of LTPP. This applies to a study by Linehan et al. (2006) which compared DBT and community treatment by experts (CTBE). Smit et al. relabelled the comparison condition of this study (CTBE) as "LTPP" by arguing that the majority of CTBE therapists in this condition described their methods as psychodynamic. However, CTBE was a mixed treatment-as-usual condition rather than an LTPP proper as some therapists also applied interpersonal therapy or humanistic therapy. Thus, results of different forms of therapy were aggregated. In addition, DBT and CTBE differed with regard to issues of treatment integrity:

1) Whereas DBT was manualized in this study, no manual was used in the CTBE condition. Moreover, the content of CTBE treatment was not prescribed in any way by the study protocol (Linehan et al., 2006, p. 759);
2) The CTBE therapists were offered supervision by a psychoanalyst, however, this does not imply that the treatments were psychodynamic, rather Linehan et al. termed the CTBE treatments "non-behavioral" (p. 763);
3) Furthermore as noted by Linehan et al. (2006, p. 759) "the CTBE therapists were not required to attend a clinical supervision group," while DBT therapists attended "weekly therapist consultation team meetings (to enhance therapist motivation and skills and to provide therapy for the therapists)."

Smit et al. included the study by McMain et al. (2009) in which DBT was compared to "general psychiatric management." Smit et al. relabelled "general psychiatric management" as "LTPP." This condition, however, was in fact general psychiatric management including a combination of "dynamically informed" therapy and symptom-targeted medication management (McMain et al., 2009, p. 1367). Furthermore, the patients assigned to this condition were not prohibited from engaging in other psychosocial treatments during the trial except for behavioral treatments (p. 1367).

Including studies in a meta-analysis which do not clearly represent the treatment under study (LTPP) can taint the Smit et al. meta-analysis as it implies poor treatment integrity in at least a subset of studies that are included.[4] Imagine the protest by proponents of CBT if a meta-anal-

4. Thus, including the Linehan et al. (2006) study is inappropriate independent of the results of the sensitivity analysis Smit et al. (2012) carried out.

ysis of CBT included studies of conditions that included and equated humanistic or interpersonal therapy with that of CBT.

(4) Different Inclusion Criteria Imply Different Research Questions

Leichsenring and Rabung (2011b) compared LTPP to shorter or less intensive forms of psychotherapy. For this meta-analytic comparison, the number of sessions conducted in LTPP was about twice as high as in the comparison conditions. The session ratio for LTPP/comparison therapy was 1.96. Compared to these shorter or less intensive forms of psychotherapy, LTPP was significantly superior in complex mental disorders (Leichsenring & Rabung, 2011b). In contrast, for the studies included by Smit et al. (2012), the session ratio was only 1.35. Accordingly, with close attention to their procedures, it becomes clear that Smit et al. compared LTPP primarily to other forms of *long-term* psychotherapy. For this reason, it is no surprise that Smit et al. did not find significant differences in outcome. In fact Smit et al. provide further evidence for the efficacy of LTPP by showing that LTPP is as efficacious as other forms of long-term therapy.

II. FURTHER ANALYSIS

(1) Recovery Rates as Primary Outcome Measure: A Reliable Measure Across Studies?

Smit et al. (2012) chose recovery rates as their primary outcome measure. The analysis of recovery rates, however, was based on only five RCTs (Dare et al., 2001; Giesen-Bloo et al., 2006; Gregory et al., 2008; Knekt et al., 2008; Svartberg, Stiles, & Seltzer, 2004). Thus, the reliability of results for their primary outcome measure is questionable. Furthermore, in each of these five studies, different definitions and measures of recovery were used. Dare et al. (2001), for example, defined recovery by an 85% increase in body weight, return of menstruation, and no bulimic symptoms, whereas Svartberg et al. (2004) used clinically significant change in self-report instruments (Jacobson & Truax, 1991). In sum, no two studies included by Smit et al. used the same definition

and measure to define recovery. Given this lack of consistency in what constituted "recovery" across studies, these results reported by Smit et al. seem to be questionable.

(2) Confusion of Post- and Follow-Up Assessments

Smit et al. (2012) included the longest follow-up data in their analysis, but did not differentiate between post-assessment and follow-up assessment data. Thus, they intermingled the results achieved post-therapy and at follow-up providing a distortion of outcome at either time point alone. In our opinion, it would be more appropriate to present effect sizes for the end of therapy and follow-up separately.

(3) Secondary Outcome Measures

Smit et al. (2012) did not find significant differences in outcome between LTPP and comparison treatments. For general psychiatric symptoms, however, Smit et al. reported an effect size of 0.69 in favor of LTPP, which is a medium to large effect size (Cohen, 1988). Although this difference did not achieve statistical significance, which may be due to the small number of studies on which this analysis was based (n = 8), it is a notable difference.

III. POTENTIAL INVESTIGATOR ALLEGIANCE EFFECTS

Researcher's own allegiances have been labelled as a "wild" card in comparisons of treatment efficacy (Luborsky et al., 1999). A recent study by Munder et al. (2012) corroborates Luborsky and colleagues' earlier study. However, investigator allegiance effects are difficult to control for as they often operate at an implicit or unconscious level. Therefore, a critical discussion of meta-analytic findings is essential (Ioannidis, 2010). Investigator allegiance effects are hard to definitively confirm, but taken together, many decisions by Smit et al. (2012) suggest the possibility of such an effect. We shall address them in the following.

(1) Involvement of Co-Authors of the Giesen-Bloo Study in the Meta-Analysis by Smit, Huibers, Ioannidis, van Dyck, van Tilburg, and Arntz

Several co-authors of the study by Giesen-Bloo et al. (2006) are among the authors of the Smit et al. (2012) meta-analysis (van Dyck, van Tilburg, and Arntz). One may wonder whether this is the reason why the study by Giesen-Bloo et al. (2006) was included in that meta-analysis; another study showing the same methodological feature (treatments not yet terminated) like that of Doering et al. (2010) was not included although it did fulfill the Smit et al. inclusion criteria.

(2) Inclusion and Exclusion of Studies

What makes the arbitrary decision to include the study by Giesen-Bloo et al. (2006) and to exclude the study by Doering et al. (2010) questionable is the fact that in all funnel plots in the Smit et al. meta-analysis the Giesen-Bloo et al. study is a clear outlier, showing larger effect sizes in favor of the comparison condition than in the other studies. Hence, including the Giesen-Bloo et al. study, and excluding the Doering et al. study which actually showed superiority of LTPP to a comparison condition, reduced the difference in effect size between LTPP and the comparison conditions. In addition, there was a public discussion whether the study by Giesen-Bloo et al. (2006) in and of itself was biased against LTPP (Leichsenring, 2009; Yeomans, 2007). In the Giesen-Bloo et al. study, competence level of LTPP therapists was considerably lower than that of schema-focused therapists. This difference in competence was not taken into account by the authors, neither with regard to the analysis of data nor in the discussion of the results. As noted above, another, possibly arbitrary, decision that raises concerns about an investigator allegiance effect can be seen in the inclusion of the studies by Linehan et al. (2006) and McMain et al. (2009) which do not represent LTPP proper. Similarly, Smit et al. (2012) excluded the study by Clarkin, Levy, Lenzenweger, and Kernberg (2007). They reported to have mailed Clarkin et al. for some data to calculate effect sizes (p. 85): "Notably, we received no answer from Clarkin et al. whom did supply data to Leichsenring and Rabung (2008; 2011)." However this is inaccurate: Clarkin and colleagues only sent Leichsenring and Rabung their manuscript

before it was published (Clarkin et al., 2007) and did not provide them with any additional data. Furthermore, Smit et al. could have contacted Leichsenring and Rabung to clarify this information and then use the existing published data similar to earlier meta-analyses (Leichsenring & Rabung, 2008, 2011b), but they did not. In addition Smit et al. did not include the complementary study by Levy et al. (2006) which provided data on attachment and self-reflective functioning. The Clarkin et al. (2007) study and the Levy et al. (2006) study showed that LTPP was as effective as DBT and superior to DBT in measures of attachment and self-reflective functioning. Thus, excluding these studies reduces the apparent effects of LTPP versus control conditions and introduces a bias to the disadvantage of LTPP.

(3) Choice of Comparison Conditions

The RCT by Dare et al. (2001) included several control conditions. Smit et al. (2012) only included the cognitive-analytic therapy condition (CAT) which actually combines cognitive and psychodynamic elements. In the Dare et al. study, however, CAT was the one comparison condition which yielded the largest effect sizes. Including only CAT as a comparison condition again introduces a disadvantage for LTPP. It is not clear why Smit et al. did not include other comparison conditions, for example family therapy or even the TAU sample reported in that study as a control condition or all comparison groups using their averaged data as previous meta-analyses have done (Leichsenring & Rabung, 2011b). Even if Smit et al. reported that including another comparison group did not change results, the arbitrary choice of CAT as the only control condition to LTPP raises the question of investigator allegiance, particularly as one of their main conclusions was that LTPP is no more effective than other control conditions, including TAU. However, when given the choice as in the study by Dare et al., they did not include TAU in the meta-analysis, rather they used another specialized active treatment.

(4) Labelling Treatment as Usual as "Straw Man"

In a subgroup analysis, Smit et al. (2012, p. 88) compared LTPP against a "straw man comparator" (a comparator without specialized psychotherapy) referring to the studies by Bateman and Fonagy (2009), Bressi, Porcellana, Marinaccio, Nocito, and Magri (2010), and Gregory

et al. (2008). However, it seems questionable to label the comparison conditions of these studies as "straw man comparators." The comparison condition of the 2009 study by Bateman and Fonagy, for example, included a manualized, structured clinical management that followed principles summarized in the NICE guidelines, and that required substantial training (NICE, 2009). Applying this same straw man criterion to studies investigating DBT, for example, would imply that most studies of DBT compared DBT against a "straw man comparator" (Leichsenring, Leibing, et al., 2011). Moreover, Smit et al. (2012, p. 81) concluded that "the recovery rate . . . was equal after LTPP or various control treatments, including treatment as usual." However, only a single treatment as usual (TAU) condition (Gregory et al., 2008) was included in this meta-analytic comparison by Smit et al. Furthermore, in the TAU condition by Gregory et al. more sessions were conducted than in the respective LTPP condition (88.7 vs. 57.7 sessions) making it harder for LTPP to be superior given the demonstration of a dose-response relationship (Gregory et al., 2008). Hence, the Smit et al. study gives the impression that LTPP is no more effective than "straw man" TAU, on the basis of a single study and not discussing session length (which they do so in other places in their paper) while such a conclusion is not warranted by their own analyses.

(5) Biased Presentation of Previous Meta-Analyses

Smit et al. (2012) discuss the results and conclusions of previous LTPP meta-analyses (de Maat et al., 2009; Leichsenring & Rabung, 2008, 2011b) by stating (p. 82): "A major objection is that both meta-analyses synthesized data from within-group differences . . . instead of between-group differences." Yet, Leichsenring and Rabung reported between-group effect sizes both immediately after the publication of their 2008 meta-analysis (Leichsenring & Rabung, 2009) as well as in another reply (Leichsenring & Rabung, 2011a) and used only between-group effect sizes in the 2011 update of their meta-analysis (Leichsenring & Rabung, 2011b).

(6) Inconsistencies Between Actual Data Examined and Conclusions

In the discussion section Smit et al. (2012, p. 98) state: "LTPP comparisons to specialized non-psychodynamic treatments, like dialectical be-

Table 1. RCTs of Long-Term Psychodynamic Psychotherapy (LTPP)

Authors	Mental Disorder	No. Patients LTPP#	Duration of LTPP	Treatment Comparison Group (No. patients)#	Duration of Comparison Treatment	RCT
Bachar et al., 1999	Eating disorders	17	12 mos, once-a-week 50-min sessions	CT (N = 17)	12 mos	yes
Bateman & Fonagy, 1999	BPD	22	18 mos, once-weekly individual therapy + thrice weekly group therapy + once-a-week expressive therapy + weekly community meeting	Psychiatric treatment as usual: inpatient treatment (90% of patients) + partial hospitalization (72%) + standard psychiatric aftercare (100%) (TAU, N = 22)	18 mos, Psychiatric review twice per mo + inpatients admission (90%) + every-2-week visits by community psychiatric nurse	yes
Bateman & Fonagy, 2009	BPD	71	92 sessions, 18 mos	SCM (N = 63)	84 sessions, 18 mos	yes
Bressi et al., 2010	Depressive or anxiety disorders	30	40 sessions, 1 yr	Treatment as usual (drug treatment + clinical interviews)	1 yr	yes
Clarkin et al., 2007; Levy et al., 2006	BPDs	LTPP1: 30, LTPP2: 30	12 mos, LTPP1: 2 individual weekly sessions; LTPP2: 1 weekly session + additional sessions as needed	DBT (N = 30)	12 mos, 1 weekly individual and group session and available telephone consultation	yes
Dare et al., 2001	Anorexia nervosa	21	24.9 sessions, 1 yr	CAT (N = 22); FT (N = 22); routine treatment (TAU, N = 19)	CAT: 12.9 sessions, 7 mos; FT: 13.6 sessions, 1 yr; TAU: 10.9 sessions, 1 yr	yes
Gregory et al., 2008	BPDs	15	57.5 sessions, 12–18 mos	TAU (N = 15)	88.7 sessions, 12–18 mos	yes

Huber & Klug, 2006; Huber et al., 2012	Depressive disorders (MDD, Recurrent Depressive Episode, or DD)	LTPP1: 35	LTPP1: 234 sessions, 39 mos	CBT (N = 34)	44 sessions, 26 mos	yes
		LTPP2: 31	LTPP2: 88 sessions, 34 mos			
Knekt et al., 2008	Depressive and anxiety disorders	128	232 sessions, 31.3 mos	STPP (N = 101); SFT (N = 97)	STPP: 18.5 sessions, 5.7 mos; SFT: 9.8 sessions, 7.5 mos	yes
Svartberg et al., 2004	Cluster C personality disorders	25	40 sessions, 16.9 mos	CT (N = 25)	40 sessions, 18.3 mos	yes
Studies Reporting Data of Ongoing Treatments						
Doering et al., 2010	BPDs	52	48.5 sessions, 1 yr	Experienced community therapists (N = 52)	18.6 sessions, 1 yr	yes
Giesen-Bloo et al., 2006	BPDs	42	231.0 sessions, 3 yrs	Schema-Focused Therapy (N = 44)	189.5 sessions, 3 yrs	yes

Note. #Intention to treat samples. BPD: Borderline Personality Disorder. CAT: Cognitive-Analytic Therapy. CBT: Cognitive Behavioral Therapy. CT: Cognitive Therapy. DBT: Dialectical Behavioral Therapy. DD: Double Depression. FT: Family Therapy. MDD: Major Depressive Disorder. SCM: Structured Clinical Management. SFT: Solution Focused Therapy. STPP: Short-Term Psychodynamic Psychotherapy. TAU: Treatment as Usual.

havior therapy and schema-focused therapy, suggest that LTPP might not be particularly effective." This conclusion is not supported by research, and particularly not by their own meta-analysis data (p. 87 and Figure 10). In fact, Smit et al. found that LTPP was as efficacious as other forms of specialized long-term psychotherapy (p. 88): "A subgroup analysis that compared LTPP against a straw-man comparator…indicated that LTPP did significantly better in the domain of target problems…but not than specialized psychotherapy treatments." Obviously, no significant differences in favor of specialized psychotherapies were found. Previous meta-analyses have not claimed that LTPP is superior to other forms of long-term psychotherapy (Leichsenring & Rabung, 2011b). Moreover, for the "specialized psychotherapies" of DBT and schema-focused therapy (SFT), only two RCTs exist which compared DBT or SFT to another specific form of therapy (Clarkin et al., 2007; Giesen-Bloo et al., 2006; Levy et al., 2006). In these RCTs, DBT and SFT were compared to LTPP. The Giesen-Bloo et al. study and its probable investigator allegiance effect was already discussed above. In the study by Clarkin et al. (2007) which was excluded by Smit et al., LTPP and DBT were equally efficacious to one another and also to supportive-dynamic psychotherapy. However, LTPP (TFP) was shown to lead to improvements in attachment and self-reflective functioning that were not found in the DBT condition (Levy et al., 2006). Thus, in this study LTPP was at least as effective and in some measures somewhat more effective compared to specialized non-psychodynamic psychotherapy. All of these results can be obtained from the original publications by Clarkin et al. (2007) and Levy et al. (2006). Thus, it is not clear why Smit et al. did not take their results into account; these studies were each published over four years ago and were widely cited.

Finally, as a point of comparison, it is worth noting that DBT, for example, has only been compared to TAU-like conditions—except for the Clarkin et al. (2007) study discussed above (Leichsenring, Leibing, et al., 2011). McMain et al. (2009) did not find DBT superior to treatments carried out by experts in clinical practice. Thus, it is in fact not clear as to whether DBT has demonstrated efficacy compared to other forms of specialized (long-term) psychotherapy.

IV. RE-EVALUATION: ANOTHER META-ANALYSIS

It is of interest to see whether the results of previous meta-analyses are stable if the studies by Giesen-Bloo et al. (2006), Bressi et al. (2010), and Doering et al. (2010) are included. For this reason, we included these studies in a new meta-analysis of RCTs on LTPP. We did this

Table 2. Comparing Long-Term Psychodynamic Psychotherapy (LTPP) with Other Forms of Psychotherapy Including the Studies by Giesen-Bloo et al. (2006) and Doering et al. (2010): Between-Group Effect Sizes (Hedges's d), Q, Ri

Outcome Domain	Number of Comparisons	Hedges's d	95% CI	Q (p)	Ri
Overall effectiveness	12	0.40	0.05–0.74	14.01 (0.23)	0.19
Target Problems	11	0.39	0.06–0.72	13.36 (0.27)	0.18
Psychiatric Symptoms	11	0.43	0.20–0.67	13.60 (0.19)	0.06
Personality Functioning	8	0.41	-0.18–1.00	6.25 (0.51)	0.36
Social Functioning	9	0.60	0.23–0.97	14.01 (0.08)	0.19

despite our judgment that the Giesen-Bloo study should be excluded from analyses for the reasons given above. Adding these three RCTs, 12 RCTs were included in the new meta-analysis (Bachar, Latzer, Kreitler, & Berry, 1999; Bateman & Fonagy, 1999, 2009; Bressi et al., 2010; Clarkin et al., 2007; Dare et al., 2001; Doering et al., 2010; Giesen-Bloo et al., 2006; Gregory et al., 2008; Huber & Klug, 2006; Huber, Zimmermann, Henrich, & Klug, 2012; Knekt et al., 2008; Levy et al., 2006; Svartberg et al., 2004). One controlled study that had been included in earlier meta-analyses on LTPP (Leichsenring & Rabung, 2008, 2011b) had to be excluded as treatment allocation was not based on randomization (Korner, Gerull, Meares, & Stevenson, 2006). The study characteristics are presented in Table 1.

The same statistical procedures were applied as in a previous meta-analysis (Leichsenring & Rabung, 2011b). Between-group effect sizes in the form of Hedges's *d* (Hedges & Olkin, 1985, p. 81) were assessed for overall effectiveness, target problems, personality functioning, and social functioning. Due to the small number of studies providing follow-up data, we assessed effect sizes only for the post-treatment data. We aggregated the effect size estimates (Hedges's *d*) across studies adopting a random-effects model which is more appropriate if the aim is to make inferences beyond the observed sample of studies (Hedges & Vevea, 1998). Tests for heterogeneity were carried out using the *Q* statistic. As the statistical power for tests of heterogeneity is often low, especially if the number of studies is small, Takkouche, Cadarso-Suarez, and Spiegelmann (1999) proposed to use random effects models and to quantify heterogeneity by means of R_i, the proportion of the total variance due to the between-study variance. R_i is assessed as R_i = tau-square/tau-square + S × var(ß), with tau = between-study variance, S = number of studies and var(ß) = total variance. Small values of R_i (< 0.40) indicate

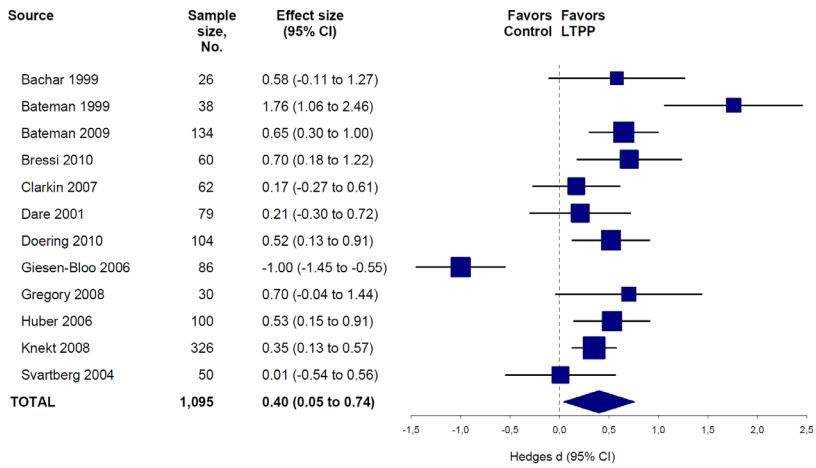

FIGURE 1. Comparative Effects of Long-Term Psychodynamic Psychotherapy on Overall Outcome

lack of heterogeneity (Etminan, Isorna, & Samii, 2005; Takkouche et al., 1999). In order to assess the degree of heterogeneity, we calculated the R_i index (Etminan et al., 2005; Takkouche et al., 1999). To obtain a mean effect sizes estimate, we used MetaWin 2.0 (Rosenberg, Adams, & Gurevitch, 1999). Statistical analyses were conducted using SPSS 15.0 and MetaWin 2.0 (Rosenberg et al., 1999; SPSS Inc., 2006).

For the 12 RCTs, the random effect model yielded the following results (Table 2). For all outcome measures, Q yielded an insignificant result for heterogeneity ($p > 0.08$). Heterogeneity as indicated by R_i was small for all measures indicating lack of heterogeneity between studies (Table 2). For overall effectiveness, target problems, general symptoms, personality functioning, and social functioning, the between-group effect sizes were 0.40, 0.39, 0.43, 0.41, and 0.60 in favor of LTPP versus comparison treatment conditions. Except for personality functioning ($p = 0.10$), all effect sizes showed significant differences in favor of LTPP (overall effectiveness $p < 0.04$, target problems $p < 0.03$, general symptoms $p < 0.01$, social functioning $p < 0.03$). As in the meta-analysis by Smit et al. (2012) the results by Giesen-Bloo et al. (2006) can be identified as outliers (Figure 1). For the 12 RCTs including the studies by Giesen-Bloo et al. (2006) and Doering et al. (2010), the session ratio LTPP/comparison condition was 1.61, below we will discuss this result in more detail.

As described above, however, we regard the inclusion of the study by Giesen-Bloo et al. (2006) as inappropriate (i.e., uncompleted treatments and limited representativeness of LTPP condition). Along with that, it

Table 3. Comparing Long-Term Psychodynamic Psychotherapy (LTPP) with Other Forms of Psychotherapy Excluding the Studies by Giesen-Bloo et al. (2006) and Doering et al. (2010): Between-Group Effect Sizes (Hedges's d), Q, Ri

Outcome Domain	Number of Comparisons	Hedges's d	95% CI	Q (p)	Ri
Overall effectiveness	10	0.51	0.25–0.77	11.63 (0.23)	0.07
Target Problems	10	0.48	0.29–0.67	9.75 (0.37)	0.01
Psychiatric Symptoms	10	0.46	0.20–0.73	12.37 (0.19)	0.07
Personality Functioning	6	0.57	0.21–0.94	5.17 (0.39)	0.04
Social Functioning	8	0.59	0.16–1.02	12.64 (0.09)	0.23

does not fulfil the inclusion criteria applied in previous meta-analyses (Leichsenring & Rabung, 2011b). In order to present valid results for the outcome of LTPP at end of treatment, we also assessed effect sizes excluding both the studies by Giesen-Bloo et al. (2006) and Doering et al. (2010) in which treatment was not terminated at the time of outcome measurement. The results are presented in Table 3.

For the remaining ten RCTs, the random effect model yielded the following results (Table 3). For all outcome measures, Q yielded an insignificant result for heterogeneity ($p > 0.08$). Again, heterogeneity as indicated by R_i was small for all measures indicating lack of heterogeneity between studies (Table 3). All between-group effect sizes were significant in favor of LTPP (overall effectiveness, $d = 0.51$, $p < 0.005$, target problems, $d = 0.48$, $p < 0.001$, general symptoms $d = 0.46$, $p < 0.01$, personality functioning, $d = 0.57$, $p < 0.009$, social functioning, $d = 0.59$, $p < 0.05$). Thus, the meta-analytic update of RCTs of LTPP confirmed the results of previous meta-analyses on LTPP (Leichsenring & Rabung, 2008, 2011b). In order to test for publication bias, we calculated correlations between sample size and between-group effect sizes across studies. None of the correlations were significant ($p > 0.31$) indicating no evidence for publication bias. This is also true if the correlation between standard error and effect sizes is used ($p > 0.15$). For the 10 RCTs the session ratio LTPP/comparison condition was 1.73. We regard these results as more valid than that of the 12-RCT meta-analysis presented above which included studies that did not fulfil our original inclusion criteria.

In sum, the meta-analytic update confirmed the results of previous meta-analyses on LTPP (Leichsenring & Rabung, 2008, 2011b), irrespective of including studies reporting data of ongoing treatments or not.

DISCUSSION

While there is a clear need for further research concerning LTPP, there is a growing body of evidence supporting the efficacy and effectiveness of LTPP in complex mental disorders. Evidence for LTPP was questioned by a recent meta-analysis by Smit et al. (2012). However, a critical review of this meta-analysis raises serious methodological concerns and the question of investigator allegiance effects. In contrast to the claims put forward by Smit et al., their results do not in any way refute or contradict that of previous meta-analyses comparing LTPP to shorter or less intensive forms of treatment (Leichsenring & Rabung, 2011b). In fact, Smit et al. compared LTPP with other forms of long-term psychotherapy—although, as noted, two studies did not meet criteria for LTPP used by Smit et al. (Dare et al., 2001; McMain et al., 2009) and another study did not represent LTPP proper (Linehan et al., 2006). We have shown that the results of Smit et al. rather suggest that various forms of long-term psychotherapy (LTPP, schema-focused therapy, DBT, CBT) are largely equivalent in terms of effect sizes, with LTPP having somewhat larger effect sizes compared to other treatments for general symptoms.

Including the studies by Giesen-Boo et al. (2006), Bressi et al. (2010), and Doering et al. (2010) did not change the results of a previous meta-analysis of LTPP (Leichsenring & Rabung, 2011b). LTPP again was significantly superior to shorter or less intensive forms of treatment in complex mental disorders. For personality functioning, the between-group effect size marginally failed to reach statistical significance (p = 0.10). This effect is somewhat smaller than in the previous meta-analysis (0.68 vs. 0.41), which may be due to the fact that the session ratio was smaller than in the previous meta-analysis (1.96 vs. 1.61). Data on dose-effect relationships suggest that a higher dose of therapy is required, especially to bring change in personality functioning (Kopta et al., 1994).

Our second (10-RCT) meta-analysis, however, provided more valid results as it did not include studies which reported data of ongoing treatments and applied our original inclusion criteria. This meta-analysis yielded effect sizes that were higher than that of the 12-RCT meta-analysis by about 0.10 for overall effectiveness, target problems, and personality functioning. This is in accordance with findings on dose-effect relationships as the session ratio here was 1.73. All between-group effect sizes were significant in favor of LTPP.

A limitation of this meta-analysis can be seen in the small number of studies that could be included and the small number of studies reporting follow-up assessments. However, a previous meta-analysis showed that effect sizes for LTPP increase significantly between post-assessment and follow-up (Leichsenring & Rabung, 2008).

A number of studies have demonstrated a fairly rapid rate of improvement early in psychotherapy for variables related to patient subjective well-being and symptomatic (acute or chronic) distress (Hilsenroth, Ackerman, & Blagys, 2001; Howard, Moras, Brill, Martinovich, & Lutz, 1996; Howard, Lueger, Maling, & Martinovich, 1993). Howard and colleagues have presented a three-phase model of recovery that progresses from subjective well-being, to symptom reduction, followed by gains in characterological, interpersonal, and social functioning (Howard et al., 1996; Howard et al., 1993).

From both a clinical and from a health-economic perspective, it is essential that future research investigates which patients may benefit sufficiently from short-term psychotherapy and which may require long-term psychotherapy. Data on dose-effect relationship suggest that most patients suffering from acute distress benefit from short-term psychotherapy (Kopta et al., 1994). For patients with chronic distress or personality disorders, about 60% and 40% of the patients recover after 25 sessions (Kopta et al., 1994). According to estimates by Perry et al., 50% of patients with personality disorder recover by 1.3 years or 92 sessions, and 75% by 2.2 years or about 216 sessions (Perry, Banon, & Floriana, 1999, p. 1318). Long-term psychotherapy, however, is associated with higher direct costs than short-term psychotherapy. For this reason, it is important to know whether the effects of long-term psychotherapy exceed those of shorter-term treatments. Results suggest that LTPP is superior to less intensive methods of psychotherapy in complex mental disorders (Leichsenring & Rabung, 2011b). For some mental disorders, for which response rates are not satisfactory, such as social anxiety disorder, experts in the field have proposed to increase treatment duration (Zaider & Heimberg, 2003). Thus, further research of long-term psychotherapy is required, not only for psychodynamic psychotherapy, but for other forms of psychotherapy as well. Long-term psychotherapy, irrespective of whether it is rooted in CBT, PDT, or another bona fide treatment approach, may be the most and perhaps the only sufficiently effective form of psychotherapy for patients with complex mental disorders.

REFERENCES

Abbass, A. A., Hancock, J., Hernderson, J., & Kisley, S. (2004). Short-term psychodynamic psychotherapies for common mental disorders. (Protocol for a Cochrane review). *The Cochrane library, Issue 2*. Chichester, UK: Wiley.

Abbass, A. A., Hancock, J. T., Hernderson, J., & Kisley, S. (2006). Short-term psychodynamic psychotherapies for common mental disorders. *Cochrane Database of Systematic Reviews, 4, CD004687*.

Bachar, E., Latzer, Y., Kreitler, S., & Berry, E. M. (1999). Empirical comparison of two psychological therapies. Self psychology and cognitive orientation in the treatment of anorexia and bulimia. *Journal of Psychotherapy Practice and Research, 8*, 115-128.

Bateman, A., & Fonagy, P. (1999). The effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *American Journal of Psychiatry, 156*, 1563-1569.

Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry*, *166*, 1355-1364.

Beck, A. T., & Rush, A. J. (1995). Cognitive therapy. In H. I. Kaplan & B. J. Sadock (Eds.), *Comprehensive textbook of psychiatry/VI* (vol. 2, pp. 1847-1856). Baltimore: Williams & Wilkins.

Bhar, S. S., & Beck, A. T. (2009). Treatment integrity of studies that compare short-term psychodynamic psychotherapy with cognitive-behavior therapy. *Clinical Psychology Science and Practice, 16*, 370-378.

Bhar, S. S., Thombs, B. D., Pignotti, M., Bassel, M., Jewett, L., Coyne, J. C., & Beck, A. T. (2010). Is longer-term psychodynamic psychotherapy more effective than shorter-term therapies? Review and critique of the evidence. *Psychotherapy and Psychosomatics, 79*, 208-216.

Bressi, C., Porcellana, M., Marinaccio, P. M., Nocito, E. P., & Magri, L. (2010). Short-term psychodynamic psychotherapy versus treatment as usual for depressive and anxiety disorders: A randomized clinical trial of efficacy. *Journal of Nervous and Mental Disease, 198*, 647-652.

Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multiwave study. *American Journal of Psychiatry, 164*, 922-928.

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

Crits-Christoph, P., & Barber, J. (2000). Long-term psychotherapy. In C. S. RE Ingram (Ed.), *Handbook of psychological change: Psychotherapy processes and practices for the 21st century* (pp. 455-473). Chichester, UK: Wiley.

Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: A meta-analysis of study quality and effect size. *Psychological Medicine, 40*, 211-223.

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

Dare, C., Eisler, I., Russel, G., Treasure, J., & Dodge, L. (2001). Psychological therapies for adults with anorexia nervosa: Randomised controlled trial of outpatient treatments. *British Journal of Psychiatry, 178*, 216-221.

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

de Maat, S., Philipszoon, F., Schoevers, R., Dekker, J., & de Jonghe, F. (2007). Costs and benefits of long-term psychoanalytic therapy: Changes in health care use and work impairment. *Harvard Review of Psychiatry*, *15*, 289-300.

Doering, S., Horz, S., Rentrop, M., Fischer-Kern, M., Schuster, P., Benecke, C., et al. (2010). Transference-focused psychotherapy v. treatment by community psychotherapists for borderline personality disorder: Randomised controlled trial. *British Journal of Psychiatry, 196,* 389-395.

Etminan, M, T. B., Isorna, F. C., & Samii, A. (2005). Risk of ischaemic stroke in people with migraine: Systematic review and meta-analysis of observational studies. *British Medical Journal*, *330,* 63-67.

Gabbard, G. (2004). *Long-term psychodynamic psychotherapy*. Washington, DC: American Psychiatric Publishing.

Gerber, A. J., Kocsis, J. H., Milrod, B. L, Roose, S. P., Barber, J. P., Thase, M. E., et al. (2011). A quality-based review of randomized controlled trials of psychodynamic psychotherapy. *American Journal of Psychiatry, 168*, 19-28.

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., et al. (2006). Outpatient psychotherapy for borderline personality disorder: Randomized trial of schema-focused therapy vs. transference-focused psychotherapy. *Archives of General Psychiatry, 63*, 649-658.

Gregory, R., Chlebowski, S., Kang, D., Remen, A. L., Soderberg, M. G., Stepkovitch, J., & Virk, S. (2008). A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder *Psychotherapy: Theory, Research, Practice, Training, 45*, 28-41.

Hedges, L., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3,* 486-504.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

Hilsenroth, M., Ackerman, S., & Blagys, M. (2001). Evaluating the phase model of change during short-term psychodynamic psychotherapy. *Psychotherapy Research, 11,* 29-47.

Howard, K., Moras, K., Brill. P., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51*, 1059-1064.

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy: Causal mediation of outcome. *Journal of Consulting and Clinical Psychology, 61*, 678-685.

Huber, D., & Klug, G. (2006). Munich Psychotherapy Study (MPS): The effectiveness of psychoanalytic long-term psychotherapy for depression. In Society for Psychotherapy Research (Ed.), *Book of abstracts. From research to practice* (pp. 154). Ulm: Ulmer Textbank.

Huber, D., Zimmermann, J., Henrich, G., & Klug, G. (2012). Comparison of cognitive-behavior therapy with psychoanalytic and psychodynamic therapy for depressed patients—A three-year follow-up study. *Zeitschrift für Psychsomatische Medizin und Psychotherapie*, *58*, 299-316.

Ioannidis, J. P. A. (2010). Meta-research: The art of getting it wrong. *Research Synthesis Methods, 1*, 169-184.

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.

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

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology, 62,* 1009-1016.

Korner, A., Gerull, F., Meares, R., & Stevenson, J. (2006). Borderline personality disorder treated with the conversational model: A replication study. *Comprehensive Psychiatry, 47,* 406-411.

Leichsenring, F. (2009). Applications of psychodynamic psychotherapy to specific disorders. Efficacy and indications. In G. Gabbard (Ed.), *The American Psychiatric Publishing texbook of psychotherapeutic treatments in psychiatry* (pp. 97-132). Washington, DC: American Psychiatric Publishing.

Leichsenring, F., Kruse, J., & Rabung, S. (in press). Efficacy of psychodynamic psychotherapy in specific mental disorders: A 2010 update. In P. Luyten, L. Mayes, P. Fonagy, M. Target., & S. J. Blatt (Eds.), *Handbook of contemporary psychodynamic approaches to psychopathology.* New York: Guilford.

Leichsenring, F., Leibing, E., Kruse, J., New, A., & Leweke, F. (2011). Borderline personality disorder. *Lancet, 377*, 74-84.

Leichsenring, F., & Rabung, S. (2008). The effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *Journal of the American Medical Association, 300,* 1551-1564.

Leichsenring, F., & Rabung, S. (2009). The effectiveness of long-term psychodynamic psychotherapy—Reply. *Journal of the American Medical Association, 301*, 932-933.

Leichsenring, F., & Rabung, S. (2011a). Double standards in psychotherapy research. *Psychotherapy and Psychosomatics, 80*, 48-51.

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

Leichsenring, F., & Rabung, S. (2012). Long-term psychodynamic psychotherapy in complex mental disorders: Update of meta-analysis—Response to Kliem et al. *British Journal of Psychiatry*, *200*, 429-430.

Leichsenring, F., Salzer, S., Hilsenroth, M., Leibing, E., Leweke, F., & Rabung, S. (2011). Treatment integrity: An unresolved issue in psychotherapy research. *Current Psychiatry Reviews, 7,* 313-321.

Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective functioning in a randomized controlled trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology, 74*, 1027-1040.

Linehan, M. M., Comtois, K. A., Murray, A. M., Brown, M. Z., Gallop, R. J., Heard, H. L., et al. (2006). Two-year randomized trial + follow-up of dialectical behavior therapy vs. therapy by experts for suicidal behaviors and borderline personality disorder. *Archives of General Psychiatry, 63*, 757-766.

Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., et al. (1999). The researcher's own allegiances: A "wild" card in comparison of treatment efficacy. *Clinical Psychology: Science and Practice, 6*, 95-106.

McMain, S., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *American Journal of Psychiatry*, *166,* 1365-1374.

Munder, T., Flückiger, C., Gerger, H., Wampold, B. E., & Barth, J. I. (2012). Is the allegiance effect an epiphenomenon of true efficacy differences between treatments? A meta-analysis. *Journal of Counseling Psychology, Epub ahead of print*.

NICE. (2009). Borderline personality disorder: Treatment and management. http://www.nice.org.uk/Guidance/CG78/NiceGuidance/pdf/English.

Perry, J., Banon, E., & Floriana, I. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry, 156*, 1312-1321.

Puschner, B., Kraft, S., Kächele, H., & Kordy, H. (2007). Course of improvement over 2 years in psychoanalytic and psychodynamic outpatient psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice, 80*, 51-68.

Rosenberg, M. S., Adams, D. C., & Gurevitch, J. (1999). *MetaWin. Statistical software for meta-analysis. Version 2.0*. Sunderland: Sinauer Associates.

Smit, Y., Huibers, M. J. H., Ioannidis, J. P. A., 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.

SPSS Inc. (2006). *SPSS 15.0 Command Syntax Reference*. Chicago, IL: SPSS Inc.

Steven, D. H. (2011). Cognitive and behavior therapy in the treatment and prevention of depression. *Depression and Anxiety, 28*, 263-266.

Svartberg, M., Stiles, T., & Seltzer, M. H. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for Cluster C personality disorders. *American Journal of Psychiatry, 161*, 810-817.

Takkouche, B., Cadarso-Suarez, C., & Spiegelmann, D. (1999). Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis. *American Jornal of Epidemiologie, 150,* 206-215.

Thoma, N. C., McKay, D., Gerber, A. J., Milrod, B. L., Ewards, A. R., & Kocsis, J. J. (2011). A quality based review of randomized controlled trials of cognitive-behavioral therapy for depression: An assessment and metaregression. *American Journal of Psychiatry*, *168*, 1-9.

Town, J., Diener, M., Abbass, A., Leichsenring, F., Driessen, E., & Rabung, S. (2012). A meta-analysis of psychodynamic psychotherapy outcomes: Evaluating the

effects of research specific procedures. *Psychotherapy (Chic), 49*(3), 276-290. doi: 10.1037/a0029564.

Yeomans, F. (2007). Questions concerning the randomized trial of schema-focused therapy vs. transference-focused psychotherapy. *Archives of General Psychiatry, 64*, 610-611.

Zaider, T., & Heimberg, R. G. (2003). Non-pharmacologic treatments for social anxiety disorder *Acta psychiatrica Scandinavica, 108 Supplementum*, 72-84.

Prof. Dr. Falk Leichsenring
Clinic of Psychosomatics and Psychotherapy
University of Giessen
Ludwigstrasse 76
35392 Giessen, Germany
Falk.Leichsenring@psycho.med.uni-giessen.de