When It Comes to Evaluating Psychodynamic Therapy, the Devil Is in the Details

Article *in* American Psychologist · February 2011
DOI: 10.1037/a0021190 · Source: PubMed

Michael D Anestis
University of Southern Mississippi

Joye C Anestis
Rutgers School of Public Health

treatment and control groups are minimal (Bhar et al., 2010; Thombs et al., 2009).

An additional criticism of the Leichsenring and Rabung (2008) review that applies to the other meta-analyses cited by Shedler (2010) relates to reliance on underpowered trials (Bhar et al., 2010; Thombs et al., 2009). Kraemer, Gardner, Brooks, and Yesavage (1998) showed that the inclusion of small, underpowered trials in meta-analyses results in substantially *overestimated* pooled effect sizes, owing to confirmatory publication bias. Statistical correction is impossible when all or most studies in a meta-analysis are underpowered. The eight studies pooled by Leichsenring and Rabung (2008), for instance, had 15–30 patients in the treatment group and power to find a moderate effect size (e.g., $\delta = 0.50$) of 0.23–0.48 (Bhar et al., 2010; Thombs et al., 2009). Effect sizes of at least 0.50–0.75 were required to achieve statistical significance in these studies. The problem is even worse than that, however, because small studies with true null effects that cross the $p < .05$ threshold do so by varying degrees. Kraemer et al. (1998) demonstrated that with 20 participants per group and a true null effect, the expected standardized effect size in a meta-analysis of statistically significant trials will be 0.90–1.00. Beyond sample size, critical assessments of the meta-analyses reviewed by Shedler (2010) must consider the poor quality of included studies, which generally failed to protect against numerous sources of potential bias, as well as the combining of overly heterogeneous trials in terms of patients treated, interventions, control groups, and outcomes (Bhar et al., 2010).

Although it may be tempting to dismiss critiques of study quality as academic, a recent study of psychotherapy for depression clearly demonstrated the danger of such an attitude (Cuijpers, van Straten, Bohlmeijer, Hollon, & Andersson, 2010). On the basis of quality criteria involving sample size considerations and standards intended to protect internal validity, Cuijpers et al. found that the effect size for high-quality studies was $d = 0.22$, compared with $d = 0.74$ for all trials. Only one psychodynamic psychotherapy trial, which had a nonstatistically significant, small effect size of $d = 0.26$, was included among studies classified as high quality. Out of all the trials reviewed in the meta-analyses cited by Shedler (2010), only one other trial (Crits-Christoph et al., 2001)—which found that standard drug counseling had a greater impact on drug use outcomes and a similar impact on associated psychological problems compared with several psychotherapy treatments, including psychodynamic psychotherapy—would have met the quality criteria proposed by Cuijpers et al. (2010).

A recent high-quality trial (Knekt et al., 2008), which was not included in any of the meta-analyses reviewed by Shedler (2010), provides an illustration of important pragmatic issues not addressed by Shedler. The trial by Knekt et al. included a comparison of the effectiveness of a mean of 232 LTPP sessions over 31 months with a mean of 9.8 sessions of nurse-administered solution-focused therapy over 7.5 months. As measured by scores on the Hamilton Depression Rating Scale, patients who received LTPP had greater depression symptom ratings at 7 months ($d = 0.21$) and 12 months ($d = 0.16$) but lower ratings at 36 months ($d = -0.26$). Results on other outcome measures were similar. All of the effect sizes for differences between groups were small and well below what is typically considered to be of clinical significance (e.g., $d = 0.50$; National Institute for Clinical Excellence, 2004). On the basis of these minimal differences, Knekt et al. concluded that LTPP provided greater benefits, and they called for more research comparing different forms of short-term versus long-term therapies. They did not note that the LTPP delivered in the study would cost approximately $29,000 to $40,600, assuming an overall cost of $125–$175 per hour, versus $735 to $980 for the nurse-delivered solution-focused therapy, assuming a cost of $75–$100 per hour.

Shedler (2010) provided an uncritical review of meta-analyses of psychodynamic psychotherapies and an unproven rationale for why critics might question evidence supporting psychodynamic psychotherapy. Rather than blanket approval or disapproval of psychodynamic psychotherapies or any other form of psychotherapy, there is a need for careful evaluation of reasonable criticisms that have been made of existing research evidence, accompanied by a consideration of pragmatic issues related to implementation and funding.

REFERENCES

Bhar, S. S., Thombs, B. D., Pignotti, M., Bassel, M., Jewett, L. R., 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 & Psychosomatics, 79,* 208–216. doi:10.1159/000313689

Crits-Christoph, P., Siqueland, L., McCalmont, E., Weiss, R. D., Gastfriend, D. R., Frank, A., . . . Thase, M. E. (2001). Impact of psychosocial treatments on associated problems of cocaine-dependent patients. *Journal of Consulting and Clinical Psychology, 69,* 825–830. doi:10.1037/0022-006X.69.5.825

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. doi:10.1017/S0033291709006114.

Knekt, P., Lindfors, O., Härkänen, T., Välikoski, M., Virtala, E., Laaksonen, M. A., . . . Helsinki Psychotherapy Study Group. (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. doi:10.1017/S003329170700164X

Kraemer, H. C., Gardner, C., Brooks, J. O., III, & Yesavage, J. A. (1998). Advantages of excluding underpowered studies in meta-analysis: Inclusionist versus exclusionist viewpoints. *Psychological Methods, 3,* 23–31. doi:10.1037/1082-989X.3.1.23

Leichsenring, F., & Rabung, S. (2008). Effectiveness of long-term psychodynamic psychotherapy: A meta-analysis. *Journal of the American Medical Association, 300,* 1551–1565. doi:10.1001/jama.300.13.1551

National Institute for Clinical Excellence. (2004). *Depression: Management of depression in primary and secondary care.* London, England: Author.

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist, 65,* 98–109. doi:10.1037/a0018378

Thombs, B. D., Bassel, M., & Jewett, L. R. (2009). Analyzing effectiveness of long-term psychodynamic psychotherapy. *Journal of the American Medical Association, 301,* 930–931. doi:10.1001/jama.2009.178

Correspondence concerning this comment should be addressed to Brett D. Thombs, McGill University, Jewish General Hospital, 4333 Cote Ste Catherine Road, Montréal, Québec H3T 1E4, Canada. E-mail: brett.thombs@mcgill.ca

DOI: 10.1037/a0021190

When It Comes to Evaluating Psychodynamic Therapy, the Devil Is in the Details

Michael D. Anestis and Joye C. Anestis
Florida State University and Psychotherapy Brown Bag

Scott O. Lilienfeld
Emory University

As Shedler (February–March 2010) noted, some researchers have reflexively and stridently dismissed psychodynamic therapy (PT) as ineffective without granting outcome studies on this modality a fair hearing. We applaud Shedler's efforts to bring

PT into the scientific mainstream and hope that his article encourages investigators to evaluate claims regarding PT's efficacy with a more objective eye. Nevertheless, as Shedler also observed, one reason for the scientific community's premature dismissal of PT is traceable to some psychodynamic practitioners' historical antipathy toward controlled research and propensity to overstate PT's efficacy. Regrettably, Shedler falls prey to the latter error by glossing over key methodological details, ignoring crucial findings that run counter to his position, and overstating the quality and quantity of the evidence base for PT. Because of space constraints, we focus only on a handful of the more serious shortcomings of Shedler's analysis (a more complete review of these issues is available from the first author on request).

To support the claim that PT boasts strong empirical support and is a worthy alternative to cognitive behavioral therapy (CBT), Shedler (2010) omitted multiple findings that contradicted his claims and discussed several problematic meta-analyses while offering minimal critical analysis of them. For example, Svartberg and Stiles (1991), not cited by Shedler, found that short-term psychodynamic psychotherapy (STPP) was superior to no treatment at posttreatment but inferior at both posttreatment and one-year follow-up to alternative approaches, including CBT and supportive psychotherapy. Svartberg and Stiles found that this difference was particularly pronounced when STPP was compared with CBT. Most notably, they observed that as the quality of the studies increased, the degree to which STPP outperformed no-treatment control groups diminished, which suggests that better-designed studies were eliminating methodological artifacts that contributed to inflated estimates of PT's efficacy. Shedler also cited a meta-analysis by Leichsenring and Leibing (2003) comparing PT with CBT for personality disorders, but he neglected to note that the overall quality of the studies in their review was low. Eight of the 11 investigations of PT were naturalistic studies rather than randomized controlled trials. These naturalistic studies used pre–post designs, which are notorious for yielding inflated effect sizes attributable to their lack of control over a plethora of methodological artifacts (e.g., spontaneous remission, regression effects, placebo effects, demand characteristics). In contrast, only three of eight studies of CBT were naturalistic, rendering the comparison between PT and CBT wildly skewed in favor of PT.

Even more important, however, *none* of the three trials in Leichsenring and Leibing's (2003) meta-analysis that compared PT with CBT unambiguously supports Shedler's (2010) conclusions. Liberman and Eckman (1981) examined individuals with a history of multiple suicide attempts (but no formal personality disorder diagnosis). Participants who received behavior therapy, considered a form of CBT in the meta-analysis, reported less suicidal ideation and fewer suicide attempts at two-year follow-up than did individuals who received "insight-oriented therapy," which was classified as a form of PT. Even though this study was underpowered ($N = 24$), it detected significant effects favoring CBT. Hardy et al. (1995) examined the treatment of depression in individuals with and without a comorbid Cluster C personality disorder. At one-year follow-up, depressed individuals with a personality disorder who received CBT did not differ from those without a personality disorder on any outcome variables. In contrast, depressed individuals with a personality disorder who received PT exhibited worse outcomes on measures of depression symptom severity, number of depression symptoms, and interpersonal problems. The third study, comparing CBT with PT for opiate addicts (Woody, McLellan, Luborsky, & O'Brien, 1985), did not report on the presence of between-groups differences; however, in an earlier article utilizing the sample from which the Woody et al. (1985) sample was drawn (Woody et al., 1983), the authors reported decidedly mixed treatment differences. Specifically, individuals who received CBT exhibited significantly greater improvement than did individuals who received PT on a range of issues related to legal problems, whereas these differences were reversed for a number of psychopathological outcome variables, including depression symptoms. The authors noted that neither therapy consistently outperformed the other in the original study. In addition, the authors did not report the degree to which individuals with diagnoses of depression and/or antisocial personality disorder responded to PT relative to CBT in the original study, thereby precluding any understanding of whether the specific symptom constellations examined in the Woody et al. (1985) study were more adequately addressed by one particular treatment approach. Moreover, the Woody et al. (1985) study compared groups ranging in size from 13 to 17 participants, rendering it severely underpowered. If we were to overlook issues of statistical power and count Woody et al.'s (1985) study as an approximate demonstration of therapeutic equivalence, this would yield a box score of two studies favoring CBT over PT and one tie. This state of affairs is hardly consistent with Shedler's sanguine conclusion regarding PT's efficacy.

Concerns regarding Shedler's (2010) review of the evidence do not end there. Shedler cited a meta-analysis by Leichsenring, Rabung, and Leibing (2004) that collapsed a wide range of diagnoses into a single category when comparing PT with CBT and other therapies. We agree with Shedler that this meta-analysis offers promising support for PT across a broad range of diagnoses. Nevertheless, the authors' methodological approach shifted attention from the clinically and theoretically important question of whether a specific treatment yielded superior results for specific diagnoses and instead investigated whether one treatment was superior to the other in general. As a consequence, it precluded an examination of whether PT was more or less efficacious for specific conditions compared with other treatments. In addition, the absence of treatment integrity checks in virtually all of the studies in this meta-analysis—a critical point not mentioned by Shedler—is a serious cause for concern and "open[s] the possibility that the absence of comparative treatment effects may be due to the manner in which the treatments were operationalized" (Bhar & Beck, 2009, p. 370). Finally, Shedler cited a meta-analysis by Leichsenring and Rabung (2008) that has also been the target of serious criticism. Follow-up analyses revealed that the vast majority of its findings were based on a computational error that resulted in the average effect size across studies exceeding the effect size in any individual study (Bhar et al., 2010).

We wholeheartedly agree with Shedler (2010) that blanket assertions that PT is worthless are scientifically unwarranted. Nevertheless, in science, a largely uncritical embrace of claims is just as problematic as is cavalier dismissal. Shedler's bold assertion that PT is as efficacious as CBT, although by no means falsified, is clearly premature given the quality and quantity of extant evidence. We encourage further investigation and evaluation of PT's efficacy, accompanied by the same dose of healthy skepticism that should apply to all psychotherapy outcome research.

REFERENCES

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. doi: 10.1111/j.1468-2850.2009.01176.x

Bhar, S. S., Thombs, B. D., Pignotti, M., Bassel, M., Jewett, L., Coyne, J. C., & Beck, A. T. (2010). Is longer term psychodynamic psy-

chotherapy more efficacious than shorter term therapies? Review and critique of the evidence. *Psychotherapy and Psychosomatics, 79,* 208–216. doi:10.1159/000313689
Hardy, G. E., Barkham, M., Shapiro, D. A., Stiles, W. B., Rees, A., & Reynolds, S. (1995). Impact of Cluster C personality disorders on outcomes of contrasting brief psychotherapies for depression. *Journal of Consulting and Clinical Psychology, 63,* 997–1004. doi:10.1037/0022-006X.63.6.997
Leichsenring, F., & Leibing, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: A meta-analysis. *American Journal of Psychiatry, 160,* 1223–1232. doi: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. doi:10.1001/jama.300.13.1551
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. doi:10.1001/archpsyc.61.12.1208
Liberman, R. P., & Eckman, T. (1981). Behavior therapy vs. insight-oriented therapy for repeated suicide attempters. *Archives of General Psychiatry, 38,* 1126–1130.
Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist, 65,* 98–109. doi:10.1037/a0018378
Svartberg, M., & Stiles, T. C. (1991). Comparative effects of short-term psychodynamic psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology, 59,* 704–714. doi:10.1037/0022-006X.59.5.704
Woody, G. E., Luborsky, L., McLellan, A. T., O'Brien, C. P., Beck, A. T., Blaine, J., . . . Hole, A. (1983). Psychotherapy for opiate addicts. *Archives of General Psychiatry, 40,* 639–645.
Woody, G. E., McLellan, A. T., Luborsky, L., & O'Brien, C. P. (1985). Sociopathy and psychotherapy outcome. *Archives of General Psychiatry, 42,* 1081–1086.

Correspondence concerning this comment should be addressed to Michael D. Anestis, Department of Psychology, Florida State University, Tallahassee, FL 32306. E-mail: anestis@psy.fsu.edu

DOI: 10.1037/a0021056

No Ownership of Common Factors

Warren W. Tryon
Fordham University

Georgiana Shick Tryon
The Graduate Center, City University of New York

Shedler's (February–March 2010) informative article "The Efficacy of Psychodynamic Psychotherapy" raised several issues worthy of comment. His choice of the word *distinctive* (p. 98) in describing aspects of psychodynamic technique is open to at least two interpretations. On the one hand, *distinctive* can have a qualitative meaning and indicate the presence of a characteristic that is not shared. For example, a sign in the Bronx Zoo distinguishes birds from all other creatures as follows: "If it has feathers it's a bird, if it doesn't, it isn't." On the other hand, *distinctive* can have a quantitative meaning and indicate that one practice has more of a common element than another practice. Careful reading of Shedler's article and the article by Blagys and Hilsenroth (2000) that forms the basis of the "seven features [that] reliably distinguished psychodynamic therapies from other therapies" (Shedler, 2010, p. 98) shows that Shedler subscribes to the latter, quantitative, definition of *distinctive*. In other words, the seven features he presented are present in both psychodynamic therapies and the cognitive-behavioral therapies to which he compares them. For example, although Shedler did not mention it, dialectical behavior therapy explicitly focuses on six of the seven features, namely, "focus on affect and expression of emotion," "exploration of attempts to avoid distressing thoughts and feelings," "identification of recurring themes and patterns," "discussion of past experience," "focus on interpersonal relations," and "focus on the therapy relationship" (Shedler, 2010, p. 99). However, in the articles that Blagys and Hilsenroth reviewed, psychodyamic therapists engaged in more of these behaviors than did cognitive-behavioral therapists.

These six features appear to us to exemplify common factors that are as basic to good psychotherapy as are Rogerian skills. Indeed, any therapist, regardless of orientation, who can engage and retain clients for up to 40 hours, which Shedler (2010) defined as short term, will necessarily have to address most, if not all, of these therapeutic elements. Although these features may have originated in the psychodynamic literature, they transcended their origins and became pantheoretical more than 30 years ago. As Bordin (1979), uncited by Shedler, indicated in his seminal article on working alliance, "The terms of the therapeutic working alliance have their origin in psychoanalytic theory, but can be stated in forms generalizable to all psychotherapies" (p. 259). In the section titled "A Rose by Another Name: Psychodynamic Process in Other Therapies," Shedler (2010, p. 103) acknowledged but did not explicitly recognize that six of his seven "distinctive" features are common factors. To imply that these factors are qualitatively unique to psychodynamic and psychoanalytic techniques is regressive. There is no ownership of common factors. Shedler's (2010) article should have been titled "Common Factors of Effective Interventions," in which case the section titled "How Effective is Psychotherapy in General?" (p. 100) could have been expanded to include the empirical evidence he presented regarding the effectiveness of modern common-factor-based psychodynamic therapy.

Shedler (2010) opined, "Undergraduate textbooks too often equate psychoanalytic or psychodynamic therapies with some of the more outlandish and inaccessible speculations made by Sigmund Freud" (p. 98). Although neither Shedler (2010) nor the several references he cited itemized these distinctive psychodynamic speculations, several come to mind: penis envy; Oedipus and Electra complexes; castration anxiety; the urethral, phallic, oral, and anal characters; along with nasal reflex neurosis and birth anxiety. If psychology textbooks have focused on these constructs, perhaps it is because they are so qualitatively distinctive. We agree with Shedler (2010) that such topics no longer deserve psychologists' attention or, hopefully, inform their clinical practice. We recommend focusing on what we share in common as a way to move our field and profession forward, and we thank Shedler (2010) for calling attention to often uncited sources of empirical support for common practices.

One can view positive treatment outcomes as validating the theoretical bases on which therapies are constructed. However, as Shedler (2010) himself pointed out, one can also be right for the wrong reasons: "The 'active ingredients' of therapy are not necessarily those presumed by the theory or treatment model" (p. 103). Sometimes treatments work for reasons other than those postulated by their authors and proponents. The efficacy of common-factor interventions supports more than psychodynamic theory. However, Shedler's (2010) focus on symptom reduction as valid outcome evidence regarding the empirical effectiveness of psychodynamic and psychoanalytic interventions is a huge departure from the psychodynamic theory of psychopathology and symptom formulation, which requires symptom substitution when underlying conflicts are not fully resolved. Tryon (2008) demonstrated that the past half century has yet to reveal any credible evidence of symptom substitution despite thousands of symptom-oriented