

What is “evidence” in psychotherapies?

The concept of evidence-based medicine (EBM), which originated in the early 1990s at McMaster University (Canada) and spread to the UK and North America, heralded an effort to place medicine on firmer scientific footing. EBM's overarching goals were twofold: to establish hortatory (“thou shall”) standards, guiding practitioners toward scientifically-supported interventions, and minatory (“thou shall not”) standards, guiding them away from scientifically-unsupported interventions.

Soon, EBM found its way into the field of psychotherapies. Evidence-based psychotherapies are commonly conceptualized as a three-legged stool. One leg comprises the best available evidence bearing on the efficacy (beneficial effects in rigorously controlled conditions) and effectiveness (beneficial effects in real-world conditions); the other two comprise clinical expertise and patient preferences/values (see Cuijpers¹ in this issue of the journal).

Still, as EBM's influence has grown, a nagging question remains: how should we conceptualize evidence in psychotherapies? Although the concept of “evidence” may seem self-explanatory, interview data suggest that academicians across multiple disciplines, including social and natural sciences, often disagree sharply regarding how to define it^{2,3}.

Probably the most influential operationalization of the evidentiary prong of the above-mentioned stool was adopted in the mid-1990s by the American Psychological Association (APA). Initially termed empirically validated therapies and later empirically supported therapies (ESTs), this prong consists of psychotherapies, typically delivered via a manual, that have been demonstrated to work for a specific psychological condition.

Modeled largely after the US Food and Drug Administration guidelines for medications, the EST criteria regard a treatment as “well-established” if it has performed better than a placebo or alternative intervention or as well as an established intervention in at least two independently conducted (performed by different research teams) randomized controlled trials or in a series of systematic within-subject studies. A secondary EST category of “probably efficacious” interventions comprises, *inter alia*, treatments that outperform a waitlist control group or that meet the aforementioned “well-established” criteria without independent replication.

Other criteria for evidence-based psychotherapies, such as the recent APA practice guidelines for post-traumatic stress disorder, depression, and childhood obesity, and those of the UK National Institute for Health and Care Excellence (NICE), consider a wider range of outcome evidence than do the EST criteria.

These organizations' laudable efforts notwithstanding, there are increasing reasons to doubt whether the current operationalization of evidence-based psychotherapies has fulfilled its mission of stemming the tide of non-scientific interventions. For example, in 2016, the US Substance Abuse and Mental Health Services Administration added thought field therapy,

which is premised on the scientifically dubious assumption that psychopathology can be treated by removing blockages in invisible and unmeasurable energy fields, to its evidence-based practice registry. In 2018, the NICE offered a “research recommendation” for a related energy therapy, emotional freedom techniques⁴. Numerous other scientifically doubtful methods, such as group drumming, equine-assisted therapy, acupuncture for depression, and music therapy for autism, have similarly claimed the evidence-based mantle. In fairness, most of these techniques might well satisfy the APA criteria for ESTs⁵.

Although a useful first step, current evidence-based guidelines, including those for ESTs, omit several key evidentiary sources needed to adequately appraise a psychotherapy's scientific grounding. To address this oversight, EST guidelines must incorporate four additional lines of evidence.

First, the replication crisis in psychology and other fields reminds us that we should be skeptical of findings unless they have been extensively replicated by multiple independent teams, ideally with offsetting theoretical allegiances. When viewed in this light, the APA EST criteria are too lax: they accord empirical support to treatments that have yielded positive findings in only two studies, and even to treatments that have yielded multiple negative findings. To enhance evidentiary rigor, the EST criteria must accommodate the full body of treatment outcome data, both positive and negative, and published and unpublished. They also must account for the methodological quality of included studies, such as sources of potential experimental bias (e.g., differential group attrition, imperfect randomization to conditions). Finally, they need to adopt statistical procedures, such as Bayesian methods, p-curve techniques, and the *r* (replicability) index, for gauging evidentiary strength and estimating publication bias⁶.

Second, evidence-based guidelines must move beyond reliance on measures of symptomatic improvement, emphasized in EST criteria, to incorporate objective and subjective criteria of everyday life functioning^{1,7}. Some patients with major depression, for example, may display significant improvement in depressive signs and symptoms (e.g., anhedonia, guilt), yet remain impaired in work and interpersonal relationships.

Third, provisional but burgeoning data from experimental and quasi-experimental studies suggest that certain treatments, such as crisis debriefing following trauma, scared straight interventions for conduct disorder, and suggestive techniques to recover ostensible memories of sexual abuse, are iatrogenic for some patients. Nevertheless, most evidence-based guidelines, including those for ESTs, overlook the possibility of harm. One challenge to addressing this omission is that many psychotherapy studies rely on unipolar outcome measures, which range from no improvement to substantial improvement; they must instead administer bipolar outcome measures, which can detect patient deterioration during and after treatment^{1,8}.

Fourth, extant evidence-based guidelines focus exclusively on outcome evidence; none considers the scientific plausibility of the treatment rationale^{5,9}. As a consequence, they open the door to all manner of pseudo-scientific interventions, many of which outperform waitlist or minimal treatment conditions. To be fair, the mode of action of many effective psychiatric interventions remains largely unknown. Yet, when interventions are premised on mechanisms that contradict well-established basic science, such as alterations in invisible energy fields, their scientific status should be suspect. Such procedures are unlikely to possess specific efficacy, that is, efficacy beyond placebo and other non-specific factors⁹.

The analysis offered here leaves unresolved the knotty question of how these diverse sources of evidence should be synthesized and weighted when appraising interventions. Reasonable arguments can be advanced for a variety of alternative evidentiary frameworks. That said, for the discipline of psychotherapy to aspire to and attain more stringent scientific standards, it must embrace a multidimensional conceptualization

of evidence, one that encompasses criteria for replicability and methodological rigor, goes beyond circumscribed indices of symptomatic improvement, accounts for potential harm, and considers all scientific evidence relevant to treatments, including basic science data bearing on treatment mechanisms.

Scott O. Lilienfeld

Department of Psychology, Emory University, Atlanta, GA, USA; School of Psychological Sciences, University of Melbourne, Melbourne, Australia

1. Cuijpers P. *World Psychiatry* 2019;18:276-85.
2. Stuart RB, Lilienfeld SO. *Am Psychol* 2007;62:615-6.
3. Scott TL, Simula BL. "You're going to have to actually think": teaching about evidence in the undergraduate curriculum. Presented at the International Institute for Qualitative Methods Conference, Brisbane, May 2019.
4. Rosen GM, Lilienfeld SO, Glasgow RE. *Lancet Psychiatry* (in press).
5. Lilienfeld SO, Lynn SJ, Bowden SC. *The Behavior Therapist* 2018;41:42-6.
6. Sakaluk JK, Williams A, Kilshaw R et al. *J Abnorm Psychol* (in press).
7. Tolin DF, McKay D, Forman EM et al. *Clin Psychol Sci Pract* 2015;22:317-38.
8. Parry GD, Crawford MJ, Duggan C. *Br J Psychiatry* 2016;208:210-2.
9. David D, Montgomery GH. *Clin Psychol Sci Pract* 2011;18:89-99.

DOI:10.1002/wps.20654