

Beware the Dodo Bird: The Dangers of Overgeneralization

Dianne L. Chambless, University of North Carolina at Chapel Hill

Luborsky et al.'s conclusion that there are no meaningful differences in the efficacy of various psychotherapies should be reconsidered for the following reasons: (a) errors in data analysis, (b) exclusion of research on many types of clients (e.g., children and adolescents), (c) faulty generalization to comparisons between therapies that have never been made, and (d) erroneous assumption that the average difference between all sorts of treatments for all sorts of problems can be assumed to represent the difference between any two types of treatment for a given problem. Concern for clients' welfare demands that psychologists be very wary of accepting the Dodo bird verdict.

Key words: psychotherapy research, empirically supported therapy, comparative efficacy of psychotherapies. [*Clin Psychol Sci Prac* 9:13–16, 2002]

Acceptance of the Dodo bird verdict is dangerous. Despite my great respect for Luborsky and his colleagues, I must disagree with their conclusion that there are no meaningful differences in outcomes of different approaches to psychotherapy (Luborsky et al., this issue). Because their assertion has profound implications for clin-

Address correspondence to Dianne L. Chambless, Department of Psychology, Room 207, Davie Hall, Cameron Avenue, UNC-CH, Chapel Hill, NC 27599-3270. E-mail: chambless@email.unc.edu.

ical practice, this is no mere academic argument. I base my arguments on both methodological and substantive grounds.

First, in the spirit of being open about allegiance, perhaps I should note that there are personal reasons I am passionate about the dangers of accepting the Dodo bird verdict. I have spent almost 30 years specializing in the treatment of people with severe anxiety disorders. Most people I see have had extensive prior therapy, and, all too often, despite their lack of response to treatment, they tell me that they were never referred for behavior therapy but found it on their own through reading articles in the popular press and the like. Many of these clients said they had voiced their discontent with their progress in therapy and asked about cognitive or behavioral treatments but were told they were resisting change, not trying hard enough, and so forth. Even sadder are the patients who were finally referred by their therapist, but only after exhausting their lifetime insurance coverage and savings. The majority (about 70%) responded to appropriate behavior therapy for their problems, but the lost years of their lives could not be recovered. At such times I am deeply distressed by the state of my profession.

On the flip side, I am also passionate about the benefit of providing the right treatment for a given client. I was a young psychotherapist when behavior therapists discovered that, even though people with agoraphobia and obsessive-compulsive disorder (OCD) had failed to benefit from systematic desensitization, they did improve rapidly with in vivo exposure (plus response prevention for OCD). With the emerging empirical evidence on flooding and in vivo exposure, our practice at the Temple University Medical School Behavior Therapy Unit, where allegiance to systematic desensitization was firm (Joseph Wolpe was the director), was transformed over a couple

of years. Concomitantly, our treatment failures became successes. This was an enormously exciting, unforgettable experience. Finally, I must own (with some pride) to having been chair of Division 12's Task Force on Promotion and Dissemination of Psychological Procedures, which introduced the focus on evidence-based psychotherapy to the United States (see Chambless et al., 1998). Having invested years in this effort, and in the process having become one of the people Dodo bird adherents most love to criticize, I might be biased by efforts to reduce any cognitive dissonance Luborsky et al.'s paper might have engendered. All that said, I still think I am right!

There are a number of problems with the meta-meta-analysis of the effects of comparative psychotherapy studies Luborsky et al. (this issue) present. First, consider the nature of the studies being compared. Here I refer to Table 1 in their article. These data are said to bear on the differences in outcome of psychotherapies, but this is not consistently the case. Take Crits-Christoph's (1991) meta-analysis of the efficacy of short-term psychodynamic psychotherapy. Two contrasts from this work are presented in Luborsky et al.'s Table 1. One concerns the comparison of dynamic therapy to nonpsychiatric treatment. This was a group of control conditions (e.g., self-help groups) that Crits-Christoph explicitly stated did not include psychotherapy. The other comparison contrasts dynamic therapy to other psychiatric treatments, a group of treatments including not only other psychotherapies but also pharmacotherapy. In addition, effect sizes from Luborsky et al.'s (1999) meta-analysis included nine comparisons of pharmacotherapy versus psychotherapy. How do any of these comparisons inform us about the equivalence of different forms of psychotherapy? There are other difficulties with Table 1, such as the inclusion in Luborsky et al.'s (1999) meta-analysis of Zitrin, Klein, and Woerner's (1978) supportive psychotherapy plus pharmacotherapy intervention as an example of dynamic therapy. The effects of drug and psychotherapy are completely confounded in this treatment condition and thus provide no information about the comparative efficacy of supportive therapy to systematic desensitization.

Second, consider the overlap among the meta-analyses presented. Reviewing the original articles, I counted 14 studies that were included in more than one of the meta-analyses in Table 1: 8 in 2 meta-analyses, 5 in 3, and 1 in 4. The approach to summarization Luborsky et al. have taken presumes that the effect sizes being aggregated are independent, but clearly they have violated this principle.

The overlap among studies sampled not only introduces statistical problems but also exaggerates the apparent stability of findings. This error is all the more surprising for these knowledgeable authors, in that they explicitly omitted 10 comparisons from Luborsky, Diguier, Luborsky, Singer, and Dickter's (1993) meta-analysis because of their overlap with the sampled studies in Luborsky et al.'s (1999) meta-analysis. Finally, in their summary of effect sizes from Svartberg and Stiles's (1991) meta-analysis of psychodynamic therapy, the authors omitted the set of studies comparing dynamic and experiential therapies. In this comparison Svartberg and Stiles found experiential therapies were significantly superior to psychodynamic therapy (recast as d , this effect size was $-.54, p < .05$).

Ignore these problems for the moment and think as if the data in Luborsky et al.'s Table 1 were correct. Does the authors' broad interpretation follow from the data? That is, are there no meaningful differences in the efficacy of various psychotherapies? I think not. First, consider what the results of this meta-meta-analysis mean.¹ At a highly aggregated level, the authors find only a small effect for differences between psychotherapies. Should this be interpreted to mean that readers can reasonably assume that whatever kind of psychotherapy they practice should be just as good as another kind for a given client? Absolutely not. Consider the following:

1. It is misleading to interpret main effects when these are modified by interactions. My reading of the literature indicates that differences among competently conducted psychotherapies may be small or nil for some problems (e.g., adult depression) but quite striking for others (e.g., agoraphobia). Even within the category of behavior therapy, clear differences emerge. For example, in four separate studies, exposure and response prevention (ERP) for OCD was compared to progressive muscle relaxation training. In all studies, ERP was significantly superior to the other treatment. As a clinician, should I conclude that I can safely ignore these findings and base my treatment on relaxation because, overall, if I were to average all possible differences between psychotherapies for all possible problems, the average difference would be small? Obviously not. (See Chambless & Ollendick [2001] for a recent review of some of the literature concerning the specificity of treatment effects.) In brief, I argue that the nature of a meta-meta-analysis that combines data on all kinds of treatments of all kinds of clients is not informative to the clinician and, in fact, is misleading.

So far I have drawn a distinction between clients with one type of disorder versus another. There is an exciting literature emerging on more subtle differences among clients and on how these interact with treatment interventions. For example, Shoham, Bootzin, Rohrbaugh, and Urry (1996) demonstrated that reactant clients were more likely to benefit from paradoxical approaches to treatment of insomnia than to standard behavioral relaxation approaches, but nonreactant clients were not. Such distinctions are missed if we conclude that all treatments produce the same results.

2. It is unwise to generalize far from the data. Luborsky et al.'s (this issue) abstract and most of their conclusions include no caveats about the limited sample of research on which they based the dodo bird finding. There are two problems here. First, it is dangerous to conclude that treatments are equally efficacious when the comparisons are lacking. To continue with my example of OCD, where are the studies that would show that interpersonal therapy, for instance, is as efficacious as exposure plus response prevention? They do not exist because only behavioral and cognitive-behavioral researchers do psychotherapy research with this population. This is probably not due to chance. Researchers do not set out to study their favored treatment with clients whom they believe will not benefit as a result, and funding agencies are unlikely to award a grant for psychotherapy research without positive pilot data showing the treatment is likely to be efficacious.

In considering the available body of literature, Luborsky et al. (this issue) limited their review to research on a narrow range of clients. Not included are studies on children and adolescents and people with psychosis, developmental disabilities, medical illness, health-endangering behaviors such as smoking, and so forth. Thus, missing are meta-analyses such as that conducted by Weisz, Weiss, Han, Granger, and Morton (1995), who found that behavioral interventions were more efficacious for children than nonbehavioral interventions. Luborsky et al. briefly mention the limitations of their sample in the discussion, but the overall thrust of the paper is to ignore this critical point.

CONCLUSIONS

It is unlikely that psychologists will ever have data available on comparisons between all major psychotherapy approaches for every common form of psychopathology.

What should a clinician do under these circumstances? I have two suggestions: First, the practitioner should examine what we do know about approaches that work for a given type of client and a given type of problem and cautiously project what those findings mean for a particular case. Second, when comparative studies are lacking, I suggest that the practitioner be guided by what we do know about the efficacy of a treatment. To continue with my OCD example, say a psychodynamic psychotherapist does an intake session with a person with OCD who wants help with this anxiety disorder. (People with OCD may, of course, seek treatment for other reasons and benefit from various approaches for differing treatment goals.) Should the practitioner ignore the literature on hundreds of clients treated successfully with ERP and use his or her preferred method of treatment, based on the rationale that no studies have been conducted to show that it is less successful than ERP? In my view, this would be unethical, at least without an informed consent process in which the clinician explains to the client that there are approaches with abundant efficacy evidence but that he or she does not propose to use them. That is, unless there is positive evidence that one treatment works as well as another (or unless there is compelling reason to do otherwise), I believe responsible practice requires using treatments for which efficacy has been demonstrated in comparisons to waiting list, attention control conditions, or alternative treatments, whenever such treatments exist. The efforts of the Division 12 Task Force (now the Committee on Science and Practice) have included assisting practitioners by easing identification of efficacious treatments (see Chambless et al., 1998 or www.wpic.pitt.edu/research/sscp/empirically_supported_treatments.htm).

To paraphrase another line from Lewis Carroll (1896/1936), I conclude that readers of Luborsky et al.'s article (this issue) should beware the dodo bird and shun overgeneralization.² A client's welfare may depend on a more cautious reading of the psychotherapy efficacy literature than we are given here.

NOTES

1. I omit comments about the authors' use of statistical corrections for allegiance and their interpretations of these findings. This issue is very important and is being addressed in the new generation of psychotherapy research in which allegiance effects are being carefully controlled, insofar as that is possible. However, the authors' arguments about this topic have been extensively discussed in a prior issue of this journal in response to

Luborsky et al.'s (1999) meta-analysis, and nothing more needs to be added here. See Hollon (1999), Jacobson (1999), Klein (1999), Lambert (1999), Shaw (1999), Shoham and Rohrbaugh (1999), and Thase (1999).

2. "Beware the Jubjub bird and shun the frumious Bander-snatch!" (from "Jabberwocky," Carroll, 1896/1936, p. 153).

REFERENCES

- Carroll, L. (1896/1936). *Jabberwocky*. In *The complete works of Lewis Carroll* (p. 153). New York: Modern Library.
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., Daiuto, A., DeRubeis, R., Detweiler, J., Haaga, D. A. F., Bennett Johnson, S., McCurry, S., Mueser, K. T., Pope, K. S., Sanderson, W. C., Shoham, V., Stickle, T., Williams, D. A., & Woody, S. R. (1998). Update on empirically validated therapies, II. *Clinical Psychologist*, *51*, 3–16.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685–716.
- Crits-Christoph, P. (1991). The efficacy of brief dynamic psychotherapy: A meta-analysis. *American Journal of Psychiatry*, *149*, 151–158.
- Hollon, S. D. (1999). Allegiance effects in treatment research: A commentary. *Clinical Psychology: Science and Practice*, *6*, 107–112.
- Jacobson, N. S. (1999). The role of the allegiance effect in psychotherapy research: Controlling and accounting for it. *Clinical Psychology: Science and Practice*, *6*, 116–119.
- Klein, D. F. (1999). Dealing with the effects of therapy allegiances. *Clinical Psychology: Science and Practice*, *6*, 124–126.
- Lambert, M. J. (1999). Are differential treatment effects inflated by researcher therapy allegiance? Could Clever Hans count? *Clinical Psychology: Science and Practice*, *6*, 127–130.
- Luborsky, L., Diguier, L., Luborsky, E., Singer, B., & Dickter, D. (1993). The efficacy of dynamic psychotherapies. Is it true that everyone has won so all shall have prizes? In N. Miller, L. Luborsky, J. P. Barber, & J. Docherty (Eds.), *Psychodynamic treatment research: A handbook for clinical practice* (pp. 497–516). New York: Basic Books.
- Luborsky, L., Diguier, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., Halperin, G., Bishop, M., Berman, J. S., & Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, *6*, 95–106.
- Shaw, B. F. (1999). How to use the allegiance effect to maximize competence and therapeutic outcomes. *Clinical Psychology: Science and Practice*, *6*, 131–132.
- Shoham, V., Bootzin, R. R., Rohrbaugh, M. H., & Urry, H. (1996). Paradoxical versus relaxation treatment for insomnia: The moderating role of reactance. *Sleep Research*, *24a*, 365.
- Shoham, V., & Rohrbaugh, M. J. (1999). Beyond allegiance to comparative outcome studies. *Clinical Psychology: Science and Practice*, *6*, 120–123.
- 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.
- Thase, M. E. (1999). What is the investigator allegiance effect and what should we do about it? *Clinical Psychology: Science and Practice*, *6*, 113–115.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, *117*, 450–468.
- Zitlin, C. M., Klein, D. F., & Woerner, M. G. (1978). Behavior therapy, supportive psychotherapy, imipramine, and phobias. *Archives of General Psychiatry*, *35*, 307–316.

Received May 1, 2001; accepted May 7, 2001.