The Dodo Bird Verdict: Status in 2014
Scott O. Lilienfeld, Emory University

The Dodo Bird, a flightless creature from the island of Mauritius, was hunted to extinction in the closing decades of the late 17th century. In contrast, the Dodo Bird verdict—named after the mythical character in Lewis Carroll’s *Alice in Wonderland* who declared (following a race) that “everyone has won, and all must have prizes”—remains alive in the early 21st century. Alert readers will note, incidentally, that I wrote “alive,” not “alive and well.” The Dodo Bird verdict, as psychotherapists and psychotherapy researchers are aware, refers to the conclusion that all psychological treatments are equal in their effects (Luborsky, Singer, & Luborsky, 1975; Seligman, 1995). Despite strenuous efforts to delegitimize the Dodo Bird verdict to the same oblivion that befell its feathered namesake (e.g., Beutler, 2002; Chambless, 2002; Chambless & Ollendick, 2001; Headsley & Dr. Giulio, 2002), some prominent authors continue to insist that the null hypothesis of therapeutic equivalence cannot be rejected (e.g., Bohart, 2000; Duncan, 2002). For example, Shedler (2010) entailed a section of his influential *American Psychologist* article “The Flight of the Dodo,” and after reviewing the early evidence for psychotherapeutic equivalence, concluded that “subsequent research has done little to alter the Dodo bird verdict” (p. 105).

In this brief and highly selective commentary, I survey recent findings bearing on the Dodo Bird verdict, and summarize this verdict’s current scientific status. I will argue that this verdict (a) is and always has been a straw person (or should I say “a straw bird?”) and (b) has for all intents and purposes been falsified, and that it is high time to consign it to the dustbin of paleontology once and for all. At the same time, important questions regarding the nature and magnitudes of specific and nonspecific effects in psychotherapy certainly remain (e.g., Wampold, 2001).

Terminological and Conceptual Confusion

Before examining the evidence bearing on the Dodo Bird verdict, we should address several terminological and conceptual issues that have contributed to all manner of confusion in the psychotherapy literature. The original meaning of the Dodo Bird verdict, introduced by psychologist Saul Rosenzweig (1956) nearly eight decades ago, did not refer to the precise equivalence of all psychotherapies. Instead, it referred to a broad equivalence in effectiveness across different “schools” (orientations) of therapy, such as psychodynamic, behavioral, and the like. Over time, however, this verdict has been magnified into the far more radical claim that all therapies are equivalent in their outcomes. In this contemporary incarnation of the Dodo Bird verdict, all or essentially all of the variance...
in therapeutic outcomes is attributable to nonspecific factors, such as the therapeutic alliance. It is worth noting, incidentally, that even the term “verdict” is a misnomer; the modern Dodo Bird is a hypothesis, albeit a dubious one.

For starters, the assertion of exact equivalence across all treatments is highly implausible on at least two grounds. First, a highly conservative and somewhat dated estimate places the total number of psychotherapies at 500 (Essler, 2000), and even this figure omits many of the more bizarre interventions, such as equine-assisted therapy for eating disorders (Christian, 2000), future life progression therapy (http://www.dheal-healing-practice.com/id32.htm), and tramiproline therapy for autism spectrum disorder (and no, I’m not making that one up; see http://www.komonews.com/news/health/tramiproline-216671201.html). The conclusion that there are no differences in outcome among any of these 500+ interventions strains credulity. Second, as many authors (e.g., Cohen, 1994; Lykken, 1968; Mehl, 1978) have noted, in the social sciences the null hypothesis of no differences across interventions is essentially always false. The null hypothesis, which is more precisely the “null hypothesis” (Cohen) of zero differences across all treatments, almost certainly cannot be true in literal form. It would imply, inter alia, that all psychological techniques are correlated zero (yes, that’s $r = .00000000000$ etc.) with the effective deployment of common factors, such as therapeutic rapport, that are themselves tied to positive client outcomes.

At this juncture, some thoughtful readers may balk. The claim of precise therapeutic outcome equivalence, they may understandably insist, must be a straw bird. Surely, no serious scholar adopts such an extreme a position, right? Au contraire. Take Duncan (2002), who noted that “the minuscule number of studies that have demonstrated superiority of one model over another” (p. 43) and even likened the comparisons among treatments to “the competition among aspirin, Advil, and Tylenol.” “All of them,” he wrote, “relieve pain and work better than no treatment at all” (p. 43). Or take Assay and Lambert (1999), who argued that “Curiously, the findings of no differences between treatments go largely unheeded” (p. 40).

Fueling the confusion, various writers have used the Dodo Bird verdict to refer to two quite different assertions, namely, (a) a main effects hypothesis or (b) an interaction hypothesis (moreover, in many cases, authors have not made clear which version of the verdict they are endorsing). The main effects hypothesis posits that collapsing across most or all psychological disorders, all psychotherapies are equal in their effects. In contrast, the interactional hypothesis proposes that there are no treatment-by-disorder statistical interactions: No therapy is preferentially better for any disorder than for any other. Such interactions, it is worth noting, provide the primary raison d’être for the impetus to develop empirically supported treatments (ESBs; Chambless & Hollon, 2001). Taken to its logical (if logical?) extreme, this interactional hypothesis implies that, conservatively speaking, the 500 (approximate number of treatments) × 300 (approximate number of DSM diagnoses) = 150,000 treatment-by-disorder interactions are precisely equal in magnitude.

Needless to say, this hypothesis is so patently absurd that we can safely reject it on a priori grounds. Does anyone seriously believe that rebirthing therapy, for example, is as effective for obsessive-compulsive disorder as it is exposure and response (ritual) prevention (ERP), or that Thought Field Therapy is as effective as applied behavior analysis for autism spectrum disorder? Even Bruce Wampold, an outspoken advocate of the position that differences among therapies are generally minimal, has been careful to note that this conclusion holds only for bona fide treatments, namely, well-established interventions that are characterized by plausible theoretical rationales (Lilienfeld, 2007; Wampold et al., 1997). Said Wampold (see DeFife, 2010) in an interview:

From my reading of the research evidence and my own research, it seems that the differences among treatments in terms of benefit to patients are small, if not negligible. This observation applies, however, to treatments that are intended to be therapeutic, as delivered by competent therapists, with a cogent psychological rationale, and contain therapeutic actions that lead to healthy and helpful changes in the patient’s life.

Regrettably, this crucial caveat appears to have been cavalierly ignored by some proponents of the Dodo Bird verdict. As a consequence, the conclusion that “the differences in outcome among therapies that have a reasonable theoretical rationale and that all work reasonably well to begin with are often minimal” has in many cases become “the differences in outcome among all therapies are minimal.” This semantic slip-page is potentially dangerous, as it can contribute to the erroneous belief that the techniques implemented by therapists are irrelevant to client outcomes.

Three Strikes Against the Dodo Bird

These key conceptual issues aside, the past decades has not been kind to the Dodo Bird. Several sources of research evidence have converged to raise serious questions regarding the blanket assertion that all treatments, even bona fide treatments, are approximately (let alone precisely) equal in their effects. I briefly summarize three of them here:

* There is growing evidence that at least some psychological treatments, such as Sealed Straight interventions for conduct disordered adolescents and critical incident debriefing to trauma-exposed individuals, can be harmful (Dimitriadou & Holon, 2010; Lilienfeld, 2007). For example, in a meta-analysis of randomized controlled trials for the prophylaxis of posttraumatic stress disorder symptoms, Litz, Gray, Bryant, and Adler (2002) found that critical incident stress debriefing displayed a slight negative effect size ($d = -1.1$) compared with no treatment or alternative treatment control conditions. Needless to say, the presence of negligible or even negative effect sizes in meta-analyses raises serious questions regarding the Dodo Bird verdict.

* In a meta-analysis of 26 randomized controlled trials ($N = 1,981$), Toliv (2010) compared cognitive-behavioral therapy (CBT) with both bona fide treatments, such as psychodynamic, interpersonal, and supportive therapies. For anxiety disorders ($d = .45$) and mood disorders ($d = .21$), but not for other conditions, CBT was significantly more efficacious than comparison interventions, with the difference attaining statistical significance for the comparison with psychodynamic treatment. CBT exerted significant effects not only on target symptoms, but also on general psychological functioning, dispelling the oft-cited claim that CBT gains often do not generalize beyond directly treated signs and symptoms (see Brewin, 1996, for a broader discussion). Although the authors of a smaller, follow-up meta-analysis (Baldessarini et al., 2013) reported no significant differences between CBT and alternative interventions, Toliv (in press) argued persuasively that their null results were a consequence of introducing excessive “noise” into the analyses. Specifically, when the analyses are
limited to studies with high methodological quality (e.g., random assignment, evaluator blinding to condition) and to global symptom measures, the clear-cut superiority of CBT over other interventions for panic disorder and generalized anxiety disorder emerges (Tolin, in press).

- Bell, Marcus, and Goodlad (2013) conducted a meta-analysis of 66 dismantling and additive studies of psychotherapy, namely, those in which full therapeutic protocols were compared with components of these protocols. Although there were no significant differences among treatments in dismantling studies, the differences among treatments in additive studies were statistically significant, albeit small in magnitude, for targeted symptoms (but not non-targeted symptoms) at termination ($d = .14$) and follow-up ($d = .25$). Although these differences are modest in size, they suggest that the addition of specific ingredients to an existing psychotherapy protocol typically yields enhanced outcomes. This finding runs counter to claims of psychotherapy outcome equivalence, which imply that only nonspecific factors are of consequence.

**Strike Four**

More recently, another strike against the Dodo Bird verdict came from a stunning study, published in *American Journal of Psychiatry*, by Paulsen et al. (2014). The authors randomized 70 patients with bulimia nervosa to either CBT or to psychoanalytic psychotherapy. Outcomes, measured using the Eating Disorder Examination interview, were assessed at 5 months and at 2-year follow-up by evaluators blind to condition assignment. Both therapies were implemented using treatment manuals developed by the study authors, with treatments delivered by well-trained therapists. Client dropout was addressed using intention-to-treat analyses.

Notably, the effects in this study were loaded heavily in favor of psychoanalytic therapy. Clients randomized to CBT received only 20 sessions of treatment over 5 months, whereas clients randomized to psychoanalytic therapy received 2 years of weekly treatment. Moreover, if any allegiance effects (see Luborsky et al., 1999) were present, they should have worked in favor of psychoanalytic therapy: the study’s 2 lead authors were proponents of this treatment and the study was carried out at a clinic that specializes in this treatment (Hollon & Wilson, 2014).

Still, the findings unambiguously favored CBT. At 5 months, 42% of bulimic patients who received CBT had ceased bingeing and purging, compared with only 6% of patients who had received psychoanalysis. At 2 years, these numbers were 45% and 16%, respectively.

Yes, this is only one study (Coyne, 2014), and we should be cautious about overhyping findings until they have been independently replicated (Pichler & Wagenmakers, 2012). At the same time, the methodological rigor of the study, conjunct with the magnitudes of the group differences, which easily surpass the hoary “interocular trauma test” of statistical significance (see Savage, 2009), should suffice to give even dedicated Dodo devotees considerable pause.

**Concluding Thoughts**

Clearly, a growing body of data indicates that previous assertions of strict equivalence across all therapeutic modalities have been essentially falsified. None of this implies, of
course, that a host of crucial questions regarding psychotherapy specificity and non-specificity do not remain to be resolved. Wampold (2001) and others may well be correct that the outcome differences among bona fide treatments have frequently been overstated, and that further attention should be accorded to the role of nonspecific effects in therapy. In particular, it will be essential to ascertain whether certain psychological conditions marked by generalized demoralization (e.g., major depression) may be responsive to a broad range of interventions, whereas conditions characterized by a less pronounced demoralization component (e.g., obsessive-compulsive disorder) may require much more targeted treatment, such as ERP. Wampold and other proponents of treatment nonspecificity have also raised constructive questions regarding the overriding emphasis on ESTs given that sizeable treatment-by-disorder interactions may be harder to come by than many of us (myself included) had once supposed. In the coming decade, a heightened emphasis on empirically supported principles of change that cut across many treatments, such as exposure, behavioral activation, and positive reinforcement of adaptive behaviors (Rosen & Davison, 2003), as well as on transdiagnostic therapeutic protocols (Barlow et al., 2010), should contribute to a thoughtful reconsideration of the relative roles of specific versus nonspecific factors in treatment processes and outcomes.

In the meantime, scholars on both sides of the debate should be able to find common ground on one central point. As my colleague Marshall Duke has noted, the Dodo Bird has become an albatross. It has increasingly impeded progress in psychotherapy research, and it has outlived its scientific utility. The verdict of strict outcome equivalence across all psychotherapies, whether in its main effect or interactional form, should at long last be declared extinct and, like its feathered counterpart, forever banished to the exhibit halls of museums.

References


Borharr, A. C. (2000). The client is the most important common factor: Clients self-healing capacities and psychotherapy. Journal of Psychotherapy Integration, 10, 127-149.


Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that “everyone has won and all must have prizes?” Archives of General Psychiatry, 32, 995-1008.


Rosen, G. M., & Davison, G. C. (2003). Psychology should use empirically supported principles of change (ESPs) and not credential trademarked therapies or other treatment packages. Behavior Modification, 27, 300-512.


Call for Applicants

The Association for Behavioral and Cognitive Therapies (ABCT, www.abct.org) invites applications for a Director of Outreach and Partnerships, with a likely start date in summer 2014. This full-time position with competitive salary and benefits reports to the Executive Director. ABCT is a 4,600 member-strong professional organization committed to the advancement of scientific approaches to the understanding and improvement of human functioning. Responsibilities include the development, implementation, and coordination of membership growth and retention strategies; building partnerships with other professional and allied organizations to advocate for and advance shared goals such as federal funding for behavioral research, recognition/funding of evidence-based approaches to prevention and treatment; and advancing ABCT’s dissemination goals. The successful candidate will be outgoing, dynamic, collaborative, and energetic; possess excellent communication skills and passion for the mission of ABCT; and be able to represent ABCT well to diverse constituents.

Required qualifications:

1. Doctoral degree in psychology or related field + ≥ five years’ experience
2. Licensed or license-eligible
3. Knowledge and passion about ABCT: cognitive-behavioral orientation and commitment to science and evidence-based principles
4. Willingness to work in NYC

Preferred qualifications:

1. Experience with professional organizations
2. Experience with outreach, membership, and marketing
3. Expertise in public policy related to behavioral health and an appreciation of the political issues and participants affecting behavioral health care
4. Experience building partnerships
5. Experience with information technology including social media
6. An academic background in population-based approaches to dissemination (e.g., MPH)

ABCT is an Equal Opportunity/Affirmative Action Employer. Send CV, contact information for five references, and a letter addressing the alignment of your experience with the qualifications to DirectorOfOutreach@abct.org. Review of applications will begin on March 24, 2014 and continue until the position is filled.