The many faces of the Dodo

An annotated bibliography of representative works on psychotherapy outcome

Brian Toomey and Bruce Ecker

The meta-analyses and large controlled trials listed below are representative of the entire history of research into psychotherapy efficacy. Taken together, they provide compelling evidence that:

1. There is no clinically significant difference in efficacy among all studied psychotherapy modalities.
2. The efficacy of each studied psychotherapy modality is equaled by the efficacy of a properly designed (structurally equivalent) placebo control.
3. When differences in efficacy of treatments and placebos have been measured, they are small and are attributable to methodological flaws, such as faulty placebo controls and investigator allegiance.
4. The efficacy of studied psychotherapy modalities is the same whether the therapist is a highly trained and experienced clinician or a minimally trained, inexperienced paraprofessional.
5. Psychotherapy and medication are equal in efficacy to each other and to psychotherapy plus medication.
6. The major classes of psychiatric drugs do not outperform properly designed placebos.
7. ECT (electroconvulsive therapy) is not more effective than placebo or “sham” ECT.

The research covered below includes data on the following treatment types:

*Psychotherapeutic*: cognitive behavioral, cognitive, systematic desensitization, behavioral, interpersonal, psychodynamic, client centered, nondirective/supportive, focusing, process experiential, gestalt, supportive, and cathartic-emotive

*Psychopharmacological*: SSRIs, MAOIs, tricyclics, benzodiazepines

*Other*: Electroconvulsive therapy (ECT)
1. Meta-analyses and direct treatment comparisons find without exception that there is no clinically significant difference in efficacy among all studied psychotherapy modalities.


Reviews 114 studies and concludes, “the evidence from these analyses supports the conjecture that the efficacy of bona fide treatments are roughly equivalent.” Differential effect sizes for various treatments were evenly distributed around zero, and “under the most liberal assumptions, the upper bound of the true effect [Cohen’s d] was about .20,” an effect size considered to be of little to no clinical significance by Cohen.


The first writer to propose factors common to seemingly different psychotherapies as an explanation for the observed equivalent outcomes of varied approaches. It was 40 years later that Rosenzweig’s ideas were empirically confirmed by Luborsky et al. (1975).


Reports a “tie score effect” summarized in the statement that “for comparisons of psychotherapy with each other, most studies found insignificant differences in proportions of patients who improved (though most patients benefited).”


Summarizes a large, non-controlled naturalistic survey of client-reported psychotherapy outcome done by the magazine Consumer Reports. The study concludes that “patients benefited very substantially from psychotherapy, that long-term treatment did considerably better than short-term treatment, and that psychotherapy alone did not differ in effectiveness from medication plus psychotherapy.” Furthermore, no specific modality of psychotherapy did better than any other for any disorder.


Reviews the literature comparing experiential and nonexperiential therapy outcome and finds that “although the difference between treatments was quite variable, the
average difference between treatments was essentially zero, echoing the oft-quoted “dodo bird” verdict. These data are thus consistent with the conclusion that experiential and nonexperiential treatments appear to be equally effective, based on the available research.” Experiential treatment subtypes included in the meta-analysis included client centered, nondirective/supportive, focusing, process experiential, gestalt, and cathartic-emotive.


Examines studies comparing cognitive-behavior therapy (CBT) and psychodynamic therapy in the treatment of personality disorder published between 1974 and 2001. The authors find no evidence that either treatment type is clinically significant superior to the other, and they emphasize the need for more studies to increase the statistical power of the conclusion.


Finds that the effects of brief dynamic psychotherapy “were about equal to those of other psychotherapies and medication” and concludes, “these data confirm previous indications that various psychotherapies do not differ in effectiveness.”


Concludes that the two treatments are equally efficacious. Previous reports of the superiority of cognitive therapy are attributed to the predominance in previous meta-analyses of studies conducted by researchers with an allegiance to the cognitive approach.


This study “examined 17 meta-analyses of comparisons of active treatments with each other... [T]he meta-analyses yielded a mean uncorrected absolute effect size for Cohen's d of .20, which is small and nonsignificant... [W]hen such differences were corrected for the therapeutic allegiance of the researchers involved in comparing the different psychotherapies, these differences tend to become even further reduced in size and significance.”

Compares outcome for patients (n=1309) who received cognitive-behavioral, person-centered, and psychodynamic therapy at one of 58 National Health Service (NHS) primary and secondary care sites in the UK. Concludes that particulars of treatment method accounted for a “comparatively tiny proportion of the variance” in outcome. The data was “consistent with previous findings that theoretically different approaches tend to have equivalent outcomes.”


Illustrates the social science application of equivalency testing, a biostatistical method often used to quantify the relative efficacy of two experimental drugs. Applied to psychotherapy efficacy studies, this method indicates the degree to which the failure to measure a difference in the efficacy of two therapies (real or placebo) can validly be interpreted as an equivalence of efficacy.

2. Standard psychotherapy modalities are not more effective then well designed, structurally equivalent placebo controls.


Reviews 21 studies in terms of the structural equivalence (degree of resemblance) between treatment and placebo controls. A placebo control was judged structurally equivalent when it matches the treatment in number and duration of sessions, group or individual format, and non-restriction of topics (for example, not merely psychoeducational). Well-designed, structurally equivalent placebo treatments were found to be as effective as the active treatments, whereas placebos lacking structural equivalence were less effective.


After reviewing all available data (22 qualifying studies) and correcting for researcher bias, the authors conclude that “when the effects of psychotherapy [for depression] were compared with those of placebo treatments, no reliable differences emerged.”

3. Methodological bias and investigator allegiance accounts for much of the reported differences between treatments

Analyzed 28 studies for investigator allegiance according to subsequent citations as well as ratings from colleagues and the authors themselves. The aggregate of the three ratings “explained 69% of the variance in outcomes”, vividly demonstrating that investigator allegiance “can distort comparative treatment results.”


Collectively these articles demonstrate that serious problems in methodological design remove credibility from the claim that a true drug effect for SSRIs is superior to the effect of a well-designed, active placebo. These problems include: 1) placebo washout strategy; 2) subjects inferring from the presence or absence of side effects whether they are in the control or treatment group and thus “penetrating the blind”; 3) unpublished studies funded by the drug industry; 4) the observed lack of correlation between outcome and drug dose; 5) investigator and publication biases.


Examines the claims that cognitive therapies are superior to other modalities and finds, contrary to other investigators who failed to correct for allegiance, no evidence of their superior efficacy after removing studies containing non-credible control treatments from the comparison pool.

### 4. Professional therapists are no more effective then paraprofessionals


Reviews 32 well-designed studies (excluding 11 due to methodological flaws) and finds no evidence for superior efficacy of extensively trained therapists versus recently or minimally trained counselors.

### 5. Psychotherapy is as effective as medication and lacks negative drug side effects


Finds that antidepressants, including tricyclics, did not yield significant differences in efficacy relative to psychotherapy, and that the combination of drug and psychotherapy was no more effective then either treatment alone.

This large study (n = 240) found cognitive therapy and medication equally effective for primary treatment of severe depression. (Faulty placebo design invalidates placebo-related findings, such as, “At 8 weeks, response rates in medications (50%) and cognitive therapy (43%) groups were both superior to the placebo (25%) group.”)


This large (n = 240), well-controlled trial found no significant evidence for difference in efficacy between the tricyclic imipramine and either cognitive-behavioral or interpersonal psychotherapy.

6. The major classes of psychiatric drugs do not outperform well-defined placebos


The authors summarize (p. 157): 1) Recent meta-analyses show that selective serotonin reuptake inhibitors (SSRIs) have no clinically significant superiority over placebo. 2) Claims that antidepressants are more effective in more severe conditions have little evidence to support them. 3) Methodological artifacts may account for the very small (and clinically insignificant) statistical margin by which SSRI efficacy exceeds that of placebos. 4) Antidepressants have not been convincingly shown to affect the long-term outcome of depression or suicide rates. 5) Given the significant doubts about their benefits and concerns about their risks, current recommendations for prescribing antidepressants should be reconsidered.


In the oldest group of community-dwelling patients to be studied to date (75 yrs plus), medication was found no more effective than placebo for the treatment of depression.


Reports a very large multi-location trial on the widely prescribed benzodiazapine, Xanax (alprazolam). After 8 weeks, initial superiority over placebos had disappeared. Also, 35% of the treatment group displayed negative withdrawal symptoms of varying severity compared to none in the placebo group. Marks et al. argue that short-term relief in a normally chronic condition does not seem to justify the adverse addiction and withdrawal effects, and that it appears likely that had the study continued and drug dependency deepened, the placebo group would soon have had better scores than the treatment group.


Found no difference in efficacy between the MAOI imipramine and the SSRI fluoxetine (Prozac).

7. **ECT is no more effective then placebo or “sham” ECT**


Summarizes: 1) Electroconvulsive therapy (ECT) is no more effective than placebo or sham ECT (a control treatment in which the patient is given anesthesia and put through shock ritual but not actually shocked) except during the period of time in which the treatment is being administered. 2) Even during while being administered, real ECT is only marginally superior to placebo. 3) Widespread claims over the past two decades that ECT is highly effective are inconsistent with the data. Notes that the transient period of marginally superior efficacy can be plausibly attributed to a masking of symptoms by ECT side effects of temporary cognitive impairment and disorientation.


On a double-blind, randomized basis, 53 patients received real ECT and 42 received sham ECT. Researchers found no difference in scores (HAM-D) at weeks 1, 12, and 28. Differences found at weeks 2 and 4 while the treatment was being administered were minor, transitory, and likely attributable to concussion-induced confusion and disorientation.