

The empirical status of cognitive-behavioral therapy: A review of meta-analyses

Andrew C. Butler^{a,*}, Jason E. Chapman^b, Evan M. Forman^c, Aaron T. Beck^a

^a *University of Pennsylvania and the Beck Institute for Cognitive Therapy and Research, United States*

^b *Medical University of South Carolina, United States*

^c *Drexel University, United States*

Received 20 September 2004; received in revised form 7 June 2005; accepted 5 July 2005

Abstract

This review summarizes the current meta-analysis literature on treatment outcomes of CBT for a wide range of psychiatric disorders. A search of the literature resulted in a total of 16 methodologically rigorous meta-analyses. Our review focuses on effect sizes that contrast outcomes for CBT with outcomes for various control groups for each disorder, which provides an overview of the effectiveness of cognitive therapy as quantified by meta-analysis. Large effect sizes were found for CBT for unipolar depression, generalized anxiety disorder, panic disorder with or without agoraphobia, social phobia, posttraumatic stress disorder, and childhood depressive and anxiety disorders. Effect sizes for CBT of marital distress, anger, childhood somatic disorders, and chronic pain were in the moderate range. CBT was somewhat superior to antidepressants in the treatment of adult depression. CBT was equally effective as behavior therapy in the treatment of adult depression and obsessive-compulsive disorder. Large uncontrolled effect sizes were found for bulimia nervosa and schizophrenia. The 16 meta-analyses we reviewed support the efficacy of CBT for many disorders. While limitations of the meta-analytic approach need to be considered in interpreting the results of this review, our findings are consistent with other review methodologies that also provide support for the efficacy CBT. © 2005 Elsevier Ltd. All rights reserved.

Cognitive-behavioral therapy is one of the most extensively researched forms of psychotherapy. Over 120 controlled clinical trials were added to the literature in the eight years between 1986 and 1993 (Hollon & Beck, 1994) and this proliferation has continued (Dobson, 2001). There are now over 325 published outcome studies on cognitive-behavioral interventions. This growth is due in part to the ongoing adaptation of CBT for an increasingly wider range of disorders and problems (Beck, 1997; Salkovskis, 1996). Yet, many questions remain regarding the overall effectiveness of CBT, its differential effectiveness by disorder, the nature of the control groups by which its effectiveness has been established, and the extent to which its effects persist following the cessation of treatment. In this paper we review evidence from meta-analyses that address these questions. Our approach is unique in that we systematically summarize findings across high-quality meta-analyses for 16 different disorders. We focus on direct comparisons of CBT to alternative treatments wherever possible.

A review of meta-analyses on CBT outcomes is particularly relevant to the ongoing debate about the comparative efficacy of different treatments (Rounsaville & Carroll, 2002). For instance, a recent review of meta-analyses and

* Corresponding author. Tel.: +1 831 372 3910; fax: +1 866 752 4401.
E-mail address: drandybutler@yahoo.com (A.C. Butler).

primary studies on CBT for depression concludes that the effectiveness of CBT has been overstated in the literature (Parker, Roy, & Eysers, 2003). The question of whether a particular psychotherapy is superior to others has produced ardent controversy. The null hypothesis, that all therapies are equally effective, has been dubbed the Dodo Bird verdict (from a line in Alice in Wonderland, “All have won and all must have prizes”). Proponents of the Dodo Bird verdict (e.g., Luborsky et al., 2002; Messer & Wampold, 2002) cite meta-analytic evidence to support their claim (“the verdict is in”). Others have taken exception, arguing that the way in which meta-analysis has been used to address this question has led to a false, or at least premature verdict (Beutler, 2002; Chambless, 2002; Howard, Krause, Saunders, & Kopta, 1997). The methodological problem at the center of this particular argument involves the practice of aggregating the outcomes for all treatments across all disorders into one meta-analysis. This approach potentially obscures real differences between specific treatments for specific disorders. The present report addresses this problem by reviewing meta-analytic evidence of the comparative efficacy of alternative treatments within disorders.

In addition to comparative efficacy, key questions exist regarding the long-term effectiveness of CBT, i.e., the extent to which intervention effects persist following the cessation of treatment. Related to this is whether these effects persist to a greater extent than do those of other treatments. Evidence suggests that the effects of psychological and especially psychopharmacological interventions substantially weaken, if not disappear entirely, once the active treatment is discontinued (e.g., Hollon, Thase, & Markowitz, 2002). However, given CBT’s focus on modifying thinking and transferring the skills learned in therapy to everyday life (i.e., making the patient his/her own therapist), treatment effects would be expected to persist following termination (Beck, 1995). Indeed, several clinical trials have pointed toward CBT’s long-term effectiveness and prevention of relapse (e.g., Simons, Levine, Lustman, & Murphy, 1984). The current review aims to determine the extent to which meta-analytic support exists for these important questions regarding the long-term effects of CBT.

Finally, two recent review articles have focused on the question of CBT’s superiority to alternative treatments for depression (Parker et al., 2003; Wampold, Minami, Baskin, & Tierney, 2002). In the present report we provide a larger context within which to interpret these reviews.

Meta-analysis has gained increasing recognition as a useful way to evaluate the efficacy of a treatment and has certain advantages as well as limitations as a review method (Cooper & Hedges, 1994; Wilkinson, 1999). In meta-analysis, treatment efficacy is quantified in terms of an effect size (ES). An ES indicates the magnitude of an observed effect in a standard unit of measurement (e.g., a standard deviation or correlation coefficient). Effect sizes have been categorized along a continuum of no effect ($ES < 0.2$), low ($0.2 \leq ES < 0.5$), medium ($0.5 \leq ES < 0.8$) and high ($ES \geq 0.8$) (Cohen, 1988). Hence, the strength of meta-analysis comes from the use of a standardized unit to compare outcomes from studies that may use different measures. Also, by averaging effect sizes across different studies and comparisons, meta-analysis increases the effective sample size and minimizes the influence of extraneous factors. This allows for a more precise evaluation of the efficacy of a treatment.

Meta-analysis is not without its limitations. One problem involves the number of comparisons included from each primary study in a meta-analysis. Some studies may report numerous comparisons and other studies few or one. Thus, findings from studies with many comparisons may be given inordinate weight in a meta-analysis. In our review we report the number of comparisons covered in each meta-analysis. Another possible limitation of meta-analysis is inattention to possible moderator variables (e.g. gender differences). When moderator variables can be identified and the sample of studies on a disorder is large enough, meta-analysis researchers may include them in their analysis. An important example of a moderator variable is the effect of researcher allegiance in outcome studies on unipolar depression. We include findings on the effect on researcher allegiance in this review. A final limitation, which also derives from inattention to study differences, is that striking successes and failures may be overlooked. Such successes and failures may be important for guiding future theoretical development and empirical study. In this review we report on emerging trends in outcome studies and suggest areas for further investigation.

1. Selection of meta-analyses

Our first objective was to identify published meta-analyses that reviewed outcome studies on cognitive-behavioral therapy. When there were multiple meta-analyses for the same disorder we selected the most extensive and methodologically rigorous meta-analysis. Typically, this was also the most recent meta-analysis. Methodological strengths we looked for were (1) inclusion of only randomized clinical trials, (2) sample-size weighting of effect sizes,

(3) analysis of heterogeneity of effect sizes and outliers, and (4) inclusion of moderator variables in analyses. These factors took precedence over the mere number of studies contained in a meta-analysis.

The identification of meta-analyses began with a search of the computer databases PsycINFO, Eric, and Medline from 1967 to July 2004 using the key terms “cognitive therapy”, “cognitive-behavioral” and “meta-analysis”. We included the term cognitive therapy since this form of treatment typically incorporates behavioral interventions. The resulting list of articles and book chapters was retrieved and reviewed. Additional meta-analyses were identified in the texts of these reviews and a second series of articles and book chapters were retrieved. The 34 reviews thus identified were then culled based on several criteria. Three meta-analyses were excluded because they were written in a language other than English. Ten were excluded because they were superseded by more extensive or methodologically rigorous reviews of the same literature. Six were excluded because they did not clearly distinguish CBT from other psychosocial interventions, one because it addressed process variables rather than outcome, and one because it did not focus on any specific disorder.

The final sample comprised 16 meta-analyses that included collectively 9995 subjects in 332 studies. These meta-analyses included 562 comparisons covering 16 disorders or populations: adult unipolar depression, adolescent unipolar depression, generalized anxiety disorder, panic disorder with or without agoraphobia, social phobia, obsessive-compulsive disorder, posttraumatic stress disorder, schizophrenia, marital distress, anger, bulimia, internalizing childhood disorders (including depressive, anxious, and somatic disorders), sexual offending, and chronic pain (not headache).

A defining feature of cognitive-behavioral therapy is the proposition that symptoms and dysfunctional behaviors are often cognitively mediated and, hence, improvement can be produced by modifying dysfunctional thinking and beliefs (Dobson & Dozois, 2001). CBT can be contrasted with purely behavioral treatments in which cognition is not considered an important explanatory variable and is not identified as a specific target for intervention. In the meta-analyses we review, there were occasionally comparisons in which studies on CBT and studies on behavior therapy were combined for analysis. When this was the case, we re-calculated the effect size for the particular comparison excluding the purely behavioral studies.

A few of the selected meta-analyses overlapped in terms of the studies reviewed. In these cases, we included the pertinent studies from only one meta-analysis. A single exception to this occurred where the same studies were examined in two different meta-analyses but with regard to different comparisons. Gould, Otto, and Pollack (1995) meta-analyzed studies comparing CT of panic disorder with alternative treatments and control groups. Oei, Llamas, and Devilly (1999) examined many of the same studies but compared patients' outcome and follow-up scores against community norms. We reviewed each meta-analysis in reference to its unique comparisons.

We have used certain conventions in reporting the findings from each meta-analysis. For instance, effect sizes are sometimes categorized as “controlled” versus “uncontrolled” (Feske & Chambless, 1995). A controlled ES expresses the magnitude of a specific treatment effect as compared to alternative treatments or control conditions. It is calculated by subtracting the post-treatment mean of the control group from the post-treatment mean of the treatment group and dividing by the standard deviation of the control group (Glass, McGraw, & Smith, 1981) or by the pooled standard deviation (Cohen, 1988). Hence, a controlled ES of 0.50 would indicate that the mean outcome score of the treatment group was one-half a standard deviation larger than the mean outcome score of the control group.

An “uncontrolled” effect size expresses the amount of improvement *within* a group from pretreatment to post-treatment. It is calculated by subtracting a group's post-treatment mean from its pretreatment mean and dividing by the pooled standard deviation. Uncontrolled effect sizes are less preferable than controlled effect sizes since they are susceptible to threats to internal validity (e.g., the effects of maturation, history, regression to the mean). Twelve of the meta-analyses included in this review report controlled sample-size weighted effect sizes. Three meta-analyses (covering obsessive-compulsive disorder, schizophrenia, and bulimia nervosa) include only uncontrolled effect sizes. These were included in this review to provide at least some indication of treatment efficacy where there was no controlled meta-analysis available for a disorder. We denote uncontrolled effect sizes as ES_u .

To facilitate interpretation, we also report U_3 scores for each comparison. U_3 represents the percentage of the scores in the lower-measured group that was exceeded by the average score in the higher-measured group (Cohen, 1988). The relationship between ES and U_3 is important to understand when interpreting the findings from meta-analyses. For example, an ES of 0.0 would indicate no treatment effect. It converts into a U_3 of 50%, which in this review would indicate that the average CT patient did as well as the average control group member. A controlled ES of 1.0

would represent a “large” treatment effect and would translate into a U_3 of 84%, indicating that the average CT patient had an outcome superior to that of 84% of the control group. Conversely, an ES of -1.0 would represent a “large” effect in the opposite direction. In this case, U_3 would again be 84% but it would indicate that the average control patient had an outcome superior to 84% of the CT group.

Meta-analytic findings were not available for some disorders. Where possible, we have included a description of the findings and conclusions of recent comprehensive qualitative review papers. However, these results should be considered preliminary, and future meta-analyses will be required to fully address the primary questions of this review.

2. Results

Table 1 lists each meta-analysis along with average effect sizes and U_3 statistics for specific comparisons. These findings are discussed below by disorder. In some cases we refer to previous meta-analyses not listed in Table 1 to illustrate important historical trends. We also refer to findings from single studies where appropriate.

2.1. Depression

CBT for depression has been studied in over 75 clinical trials since 1977 (Gloaguen, Cottraux, Cucherat, & Blackburn, 1998). Parker et al. (2003) reviewed meta-analyses and primary studies in this area and concluded that CBT is not as effective as proponents of the therapy suggest. However, they excluded from their review high-quality clinical trials showing cognitive-behavioral therapy to be superior to alternative treatments at follow-up (see Shapiro et al., 1994 for a comparison with psychodynamic therapy; Blatt, Zuroff, Bondi, & Sanislow, 2000 for a comparison with antidepressant medication). Since Parker et al. did not report the criteria by which they selected research studies for their review, it is difficult to interpret their conclusions. Dobson (1989) conducted the first meta-analysis on this topic and found CT to be superior to untreated controls, wait list, pharmacotherapy, behavior therapy, and a heterogeneous group other therapies. Gaffan, Tsaousis, and Kemp-Wheeler (1995) hypothesized that Dobson’s findings were positively biased toward CBT by researcher allegiance; the tendency for authors to prefer one therapy to another. They reanalyzed the same set of studies reviewed by Dobson while controlling for researcher allegiance and found that researcher allegiance accounted for half the difference between CBT and other treatments. They also analyzed a later set of 37 studies on CBT for depression and found no evidence of researcher allegiance. A possible explanation for the difference in findings is that researcher allegiance influences study outcomes when a new treatment is first tested, but its influence fades over time. In the second analysis, CBT was still found to be superior to waiting list, attentional controls, and a group of other various psychotherapies.

Gloaguen et al. (1998) conducted the most extensive and methodologically rigorous meta-analysis on CBT for depression to date. Their findings are displayed at the top of Table 1. CBT was found to be superior when compared with waiting list or placebo controls. Since waiting list and placebo conditions are not equivalent, future meta-analyses should evaluate comparisons of CBT with each of these conditions separately (Parker et al., 2003). The difference between CBT and behavior therapy was not statistically significant. Although some of the studies comparing CBT with behavior therapy were done with non-clinical samples, findings for clinical and non-clinical samples led to the same conclusion of treatment equivalence.

Gloaguen et al. (1998) also compared CBT to a group of therapies categorized as miscellaneous and found that CBT was modestly superior (ES=0.24). Wampold et al. (2002) pointed out that some of the therapies in the Gloaguen et al.’s miscellaneous category (e.g., progressive muscular relaxation) do not represent comprehensive treatments for depression. Based on criteria that they developed, Wampold et al. further categorized these miscellaneous therapies into bona fide and non-bona fide therapies for depression. They then compared CBT to the group of therapies they identified as bona fide and found that CBT was only marginally superior (ES=0.16, $p < .06$).

A comparison of CBT (or any psychotherapy) with a control group consisting of a heterogeneous collection of miscellaneous therapies, as done by Gloaguen et al., can at best provide limited information. Future meta-analyses are needed that provide direct comparisons between CBT and specific alternative therapies. Due to the large number of practitioners of psychodynamic therapy, comparisons between CBT and psychodynamic therapy for

Table 1
Meta-analyses of cognitive therapy outcomes

Meta analysis	Comparisons	Outcome variable	Comparison group	ES	U_3 (%)
<i>Adult unipolar depression</i>					
Gloaguen et al. (1998)	20	Beck depression inventory	Wait list or placebo	0.82	79
	17	"	Antidepressants	0.38	65
	13	"	Behavior therapy	0.05	52
	22	"	Other miscellaneous psychotherapies	0.24	59
<i>Adolescent unipolar depression</i>					
Reinecke et al. (1998)	5	Within-study average of various depression measures	Wait list	1.11	87
	1	Mood and feelings questionnaire	Relaxation	0.75	77
	1	Beck depression inventory	Supportive therapy	0.55	71
	4	Within-study average of various depression measures	Wait list at follow-up (mean follow-up interval=6.3 weeks)	0.55	71
	1	Mood and Feelings Questionnaire	Relaxation at follow-up (3 months)	0.45	67
	1	Beck depression inventory	Supportive therapy at follow-up (3 months)	1.12	87
<i>Generalized anxiety disorder</i>					
Gould, Otto et al. (1997)	4	Various self-report measures of anxiety or worry	Wait list or no treatment	0.82	79
	1	"	Nondirective therapy	0.69	75
	1	"	CT+relaxation vs. nondirective+relaxation	0.72	76
	2	"	Pill placebo	1.26	90
<i>Panic disorder with agoraphobia</i>					
Oei et al. (1999)	7	Fear questionnaire total score	Community norms	-0.48 ^a	69
	11	Fear questionnaire agoraphobia score	Community norms	0.17 ^a	57
	7	Fear questionnaire total score	Community norms—Follow-up (interval unspecified)	-0.47 ^a	68
	11	Fear questionnaire agoraphobia score	Community norms — Follow-up (interval unspecified)	0.24 ^a	59
<i>Panic disorder with or without agoraphobia</i>					
Gould et al. (1995)	3	Within-study average of various panic measures	Cognitive interventions+situational exposure vs. situational exposure only	0.53	70
	7	"	Cognitive restructuring+interoceptive exposure vs. wait list	0.91	82
	1	"	Cognitive restructuring+interoceptive exposure vs. pill placebo	0.65	74
	1	"	Cognitive restructuring+interoceptive exposure vs. psychological placebo	1.29	90
	2	"	Cognitive restructuring alone vs. wait list or minimal contact control	0.75	77
	1	"	Cognitive restructuring alone vs. situational exposure alone	-0.95	83
	5	"	"Slippage" in effect size from post-treatment to follow-up (mean follow-up interval=12.6 months)	0.02	52
<i>Social phobia</i>					
Gould, Buckminster et al. (1997)	7	Within-study average of various social phobia measures	Wait list or placebo attention controls	0.93	82
<i>Obsessive-compulsive disorder</i>					
Van Balkom et al. (1994)	4	Obsessive-compulsive symptoms: patient rating	Reduction from pre- to post-treatment	1.30 _u	90

(continued on next page)

Table 1 (continued)

Meta analysis	Comparisons	Outcome variable	Comparison group	ES	U_3 (%)
<i>Obsessive-compulsive disorder</i>					
Van Balkom et al. (1994)	2	Obsessive-compulsive symptoms: assessor rating	Reduction from pre- to post-treatment	1.86 _u	97
Abramowitz (1997)	4	Behavioral measures and Yale-Brown obsessive-compulsive scale	Exposure with response prevention	0.19 _u	58
<i>Posttraumatic stress disorder</i>					
National Collaborating Centre for Mental Health (2005)	14	Likelihood of having a PTSD diagnosis after treatment	Wait list	0.47 ^b	
	8	Severity of PTSD symptoms (self-report measures)	Wait list	1.70	96
	13	Severity of PTSD symptoms (clinician-rated measures)	Wait list	1.36	91
	13	Depression symptoms	Wait list	1.20	89
	10	Anxiety symptoms (self-report measures)	Wait list	0.94	83
	6	Likelihood of having a PTSD diagnosis after treatment	Eye Movement Desensitization and Reprocessing (EMDR)	1.03 ^b	
	6	Severity of PTSD symptoms (self-report measures)	EMDR	0.31	62
	6	Likelihood of having a PTSD diagnosis after treatment	Stress management therapies (not trauma focused)	0.78 ^b	
	3	Severity of PTSD symptoms (self-report measures)	Stress management therapies (not trauma focused)	0.37	65
	5	Likelihood of having a PTSD diagnosis after treatment	Other therapies	0.71 ^b	
	3	Severity of PTSD symptoms (self-report measures)	Other therapies	1.18	88
<i>Schizophrenia</i>					
Rector and Beck (2001)	4	Positive symptoms (Various clinician rating scales)	Reduction from pre- to post-treatment	1.28 _u	90
	3	Negative symptoms (various clinician rating scales)	Reduction from pre- to post-treatment	0.91 _u	82
	2	Total symptoms (various clinician rating scales)	Reduction from pre- to post-treatment	1.50 _u	93
Dunn and Schwebel (1995)	3	Various questionnaire measures of marital adjustment	Untreated control couples at end of treatment	0.71	76
	2	"	Untreated control couples at follow-up (mean interval=9.2 months)	0.54	71
<i>Anger</i>					
Beck and Fernandez (1998)	50	Various measures of anger and aggression	No treatment	0.70	76
<i>Bulimia nervosa</i>					
Whittal et al. (1999)	17	Binge frequency	Reduction from pre- to post-treatment	1.28 _u	90
	24	Purge frequency	Reduction from pre- to post-treatment	1.22 _u	89
	13	Various self-report measures of eating attitudes	Improvement from pre- to post-treatment	1.35 _u	91
<i>Internalizing childhood disorders</i>					
Grossman and Hughes (1992)	4	Various measures of anxiety	No treatment or psychological placebo	0.93	82
	9	Various measures of depression	No treatment or psychological placebo	0.87	81
	7	Various measures of somatic disorders	No treatment or psychological placebo	0.47	68

Table 1 (continued)

Meta analysis	Comparisons	Outcome variable	Comparison group	ES	U_3 (%)
<i>Sexual offending</i>					
Nagayama Hall (1995)	5	Recidivism (mean follow-up interval=4.5 years)	No treatment	0.35	64
<i>Chronic pain (not headache)</i>					
Morley et al. (1999)	16	Pain experience	Wait list	0.33	63
	22	Mood	Wait list	0.40	66
	16	Cognitive coping and appraisal	Wait list	0.50	69
	5	Pain expression behavior	Wait list	0.49	69
	7	Activity level	Wait list	0.48	69
	15	Social role functioning	Wait list	0.61	73
	17	Pain experience	Heterogeneous collection of treatments (not including behavior therapy or biofeedback)	0.26	60
	24	Mood		-0.06	53
	23	Cognitive coping and appraisal		0.33	63
	8	Pain expression behavior		0.31	62
10	Social role functioning		0.10	54	

Positive ES scores indicate superiority of cognitive therapy over the comparison group. Negative ES scores indicate superiority of the comparison group over cognitive therapy. U_3 represents the percentage of subjects in the inferior group whose outcome scores were exceeded by the average subject in the superior group. Subscript 'u' denotes an uncontrolled effect size.

^a In these analyses only, a negative effect size indicates symptom scores below community norms. A positive effect size indicates symptom scores above community norms.

^b Relative risk (RR) is used as the index of effect size in these comparisons. For dichotomous outcomes, RR is the ratio of the treatment event to the control event. An RR of 1 indicates no difference between treatment and control conditions. An RR of 0.75 indicates that the event rate associated with CT is three-quarters of that with the control condition, or in other words, the relative risk reduction is 25. No U statistic is reported for these effect sizes since RR is as intuitively interpretable as U .

depression are of particular interest. There are relatively few direct comparisons and not enough to provide the basis for a meaningful meta-analysis. Two large well-controlled trials have been conducted. The first was the Treatment of Depression Collaborative Research Project (TDCRP) conducted in the 1980s. This study compared the efficacy of CBT, interpersonal psychotherapy, and pharmacotherapy for depression. Outcomes for CBT and interpersonal therapy were generally equivalent (Elkin et al., 1989), although CBT did not fare as well as IPT or antidepressant medication among the more severely depressed patients in this study. In the 1990s, Shapiro et al. conducted a major study comparing the efficacy of CBT and psychodynamic-interpersonal (PIT) for depression. This study also experimentally manipulated the duration of treatment (8 sessions versus 16 sessions). At the end of treatment CBT and PIT were equally effective (Shapiro et al., 1994). However, 1 year after completing treatment, patients who had received 8 sessions of PIT did worse on almost all measures compared with the other three treatment conditions (8-session CBT, 16-session PIT, and 16-session CBT).

The intervention for depression that is most widely used in the United States is antidepressant medication (Olfson & Klerman, 1993). Gloaguen et al. (1998) found that CBT was significantly better than antidepressant medication. This conclusion may overstate the case for CBT since Gloaguen et al. (1998) included some early studies comparing CT with medications, which had methodological features that favored CT. One study (Rush, Beck, Kovacs, & Hollon, 1977) started tapering medications 1 week prior to the end of therapy, and another study (Blackburn, Bishop, Glen, Whalley, & Christie, 1981) was an effectiveness study in which the comparison group showed an unusually low response to medication. More recently, however, a major high-quality controlled trial comparing CT with a commonly prescribed serotonin reuptake inhibitor (paroxetine) found that cognitive therapy was equally effective for the initial treatment of moderate to severe major depression (DeRubeis et al., 2005).

The combination of cognitive and pharmacotherapy for depression is widely practiced in the community. In their narrative review, Hollon and Beck (1994) reported that the combination of CT and pharmacotherapy has typically been associated with a small (0.25 standard deviation) advantage over either modality alone. There is some evidence that the combination of psychotherapy and antidepressant medication leads to significantly better outcomes with severely depressed patients (Thase et al., 1997). However, the Thase et al. analysis aggregated several forms of

psychotherapy. The question of whether adding medication to CBT for severely depressed patients leads to significantly better outcomes has not yet been the subject of a meta-analysis.

Given the potentially marginal differences in outcome between single versus combination treatments, and the higher relapse rates with pharmacotherapy, long-term cost-effectiveness becomes a key issue. In a cost-effectiveness analysis of several outcome studies on depression, Antonuccio, Thomas, and Danton (1997) found that over a 2-year period, fluoxetine alone resulted in 33% higher expected costs than individual cognitive behavior therapy and the combination treatment resulted in 23% higher costs than cognitive behavior therapy alone.

Reinecke, Ryan, and DuBois (1998) conducted a meta-analysis of CT of depressive symptoms in adolescents. There are far fewer studies of CT with this population and most studies have rather small sample sizes. The reported effect sizes do suggest that CT is superior to wait list, relaxation therapy, and supportive therapy at end of treatment. However, conclusions from this meta-analysis should be considered preliminary. This is especially true given that a recent multi-site controlled trial found that CBT was equivalent to pill-placebo for reduction of depressive symptoms in adolescents (Treatment for Adolescents With Depression Study (TADS) Team, US, 2004).

Presented next are the follow-up findings, where available, for the studies described above. The effectiveness of CBT in preventing relapse of depression was evaluated by Gloaguen et al. (1998). In their meta-analysis, they examined a subset of eight studies that compared relapse rates for CBT versus antidepressants at least a year after discontinuation of treatment. They concluded that CBT was associated with a “preventative” effect in five of the eight studies. On average, only 29.5% of CBT patients relapsed versus 60% of patients treated with antidepressants. In the treatment of depressed adolescents, Reinecke et al. (1998) found that the superiority of CT to wait list, relaxation, and supportive therapy was maintained at 6–12 weeks follow-up.

Additionally, to augment these findings, we have included follow-up findings from recent comprehensive reviews and large-scale trials. DeRubeis and Crits-Cristoph (1998), in their review of treatment efficacy, obtained highly similar figures by averaging 1-year relapse rates across three large-scale clinical trials, finding a substantial difference between CT (26%) and pharmacotherapy (64%). Paykel et al. (1999) examined whether 16 sessions of CBT would help prevent relapse among depressed patients who had responded partially to pharmacotherapy. They found that the cumulative relapse rate at 68 weeks was significantly lower for antidepressant plus CBT patients. Recently, a large multi-site clinical trial found that cognitive therapy had an enduring effect for moderate-to-severely depressed patients that extended beyond the end of treatment and was equivalent to the effect of keeping patients on antidepressant medication (see Hollon et al., 2005). In this study, patients who responded to cognitive therapy in a randomized controlled trial (DeRubeis et al., 2005) were withdrawn from treatment and compared during a 12-month period with medication responders who had been randomly assigned to either continuation medication or placebo withdrawal. Patients withdrawn from CT were significantly less likely to relapse during the 12-month follow-up period than patients withdrawn from medications (31% versus 76%), and no more likely to relapse than patients who kept taking continuation medication (31% versus 47%).

2.2. Generalized anxiety disorder

The effectiveness of CT for generalized anxiety disorder was recently investigated in a meta-analysis by Gould, Otto, Pollack, and Yap (1997). Again, the number of studies is small and future replication with more studies is needed. The preliminary results are strong, with CT being substantially superior to wait list or no-treatment controls, nondirective therapy, and pill placebo. Gould et al. reported that CT and pharmacotherapy showed similar levels of improvement from pre- to post-treatment ($ES_u = 0.70$ and 0.61 , respectively). Benzodiazepines ($ES_u = 0.70$) were the most commonly evaluated medications in these clinical trials.

With respect to long-term outcome, Gould et al. (Gould, Buckminster, Pollack, Otto, & Yap et al., 1997; Gould, Otto et al., 1997) found that CBT treatment effects were maintained through at least 6 months post-treatment ($ES = 0.05$). These results were not available for pharmacotherapy due to a lack of long-term outcome data. DeRubeis and Crits-Cristoph's (1998) review of GAD clinical trials found a robust persistence of effects for CT over 3, 6 and 12 months. Additionally recent reports indicate a persistence of CT effects as far as 8–10 years post-treatment (Durham, Chambers, MacDonald, Power, & Major, 2003). Moreover, in general, CT has demonstrated stronger long-term effects relative to other treatments. One trial, for example, reported that 58% of CT evidenced clinically significant change at 12-month follow-up, compared with 38% for applied relaxation (Borkovec & Costello, 1993).

2.3. Panic disorder

The outcome findings on CT for panic disorder are supportive and robust. Gould et al. (1995) conducted a meta-analysis comparing cognitive interventions with exposure treatments and various control conditions. Among the cognitive-behavioral treatments, those that combined cognitive restructuring with interoceptive exposure showed the strongest effect ($ES=0.88$). This combination of interventions is the standard CBT approach to treating panic disorder. Cognitive-behavioral treatments in general had the highest effect size ($ES=0.68$) compared to pharmacological treatments ($ES=0.47$) or combination treatments ($ES=0.56$). Perhaps most impressive is that the CBT treatment showed virtually no “slippage” in effect size, (-0.07) by 1-year follow-up as compared to sizable slippage (-0.46) for pharmacological treatment. Importantly, there was also a significantly lower attrition rate among cognitive-behavioral treatments (5.6%) compared to pharmacological treatments (19.8%).

Oei et al. (1999) compared panic patients' scores at the end of treatment and at follow-up against community norms. They computed effect sizes by subtracting the treatment group mean from the community mean and then dividing by the community standard deviation. They found that CT reduced symptoms to levels near or below those found in the general population by end of treatment and these treatment gains were maintained over an unspecified follow-up interval.

2.4. Social phobia

Cognitive-behavioral treatments for social phobia have been subjected to meta-analysis twice in recent years (Feske & Chambless, 1995; Gould, Buckminster et al., 1997). The most recent and extensive meta-analysis by Gould et al. replicated the findings of Feske and Chambless showing that CT was superior to wait-list and placebo attention controls ($ES=0.93$), and similar to exposure interventions without cognitive restructuring ($ES=0.89$) and the combination of the two interventions ($ES=0.80$). They found that the average ES for pharmacological interventions was 0.62, with selective serotonin reuptake inhibitors ($ES=1.89$) superior to benzodiazepines ($ES=0.72$). Neither meta-analysis reports on long-term effects of social phobia treatments. However, DeRubeis and Crits-Cristoph (1998) report that gains were maintained over the long-term in five out of the five large clinical trials reviewed.

2.5. Obsessive-compulsive disorder

Exposure with response prevention (ERP) is sometimes considered the psychological treatment of choice for obsessive-compulsive disorder, but there is growing evidence that CBT is equally effective (Abramowitz, 1997). A meta-analysis conducted by Van Balkom et al. (1994) found that cognitive or cognitive-behavioral treatments for OCD led to substantial reductions in OCD symptoms as rated by patients ($ES_u=1.30$) and clinical assessors ($ES_u=1.86$). Furthermore, these treatment effects tended to persist at 6-month and even 12-month follow-ups. Abramowitz (1997) found that the effect sizes of the four direct comparisons of cognitive-behavior therapy with ERP suggest no difference between the treatments. More recently, in a controlled clinical trial, a combined treatment (cognitive therapy plus ERP) was found to confer clinical gains over and above ERP alone in the treatment of medication-refractory OCD (Rector et al., 2005).

2.6. Posttraumatic stress disorder

The National Collaborating Centre for Mental Health in Great Britain recently completed a rigorous meta-analytic review of the clinical research literature on PTSD. Commissioned by the National Institute for Clinical Excellence (NICE), this meta-analysis served as the basis for the National Clinical Practice Guideline on PTSD (Royal College of Psychiatrists & The British Psychological Society, 2005). The studies reviewed on trauma-based CBT covered a range of patient populations including survivors of accidents, assault, sexual assault (including childhood sexual assault), domestic violence, military combat, ‘mixed’ trauma groups and refugees with multiple traumatic events. Overall, the findings showed clinically important benefits for trauma-focused CBT over wait list control groups on all measures of PTSD symptoms (mean $ES=1.49$). There was also evidence that trauma-focused CBT has clinically significant effects on depression and anxiety in these patients. The other treatment supported by the review was eye-movement

desensitization and reprocessing (EMDR). The reviewers note that both trauma-focused CBT and EMDR specifically address the PTSD sufferers' troubling memories of a traumatic event and the personal meanings of the event and its consequences. The meta-analysis also found that direct comparisons of these two approaches did not reveal any significant advantages for one over the other, with respect to either treatment outcome or the speed of therapeutic change.

2.7. Schizophrenia

While pharmacotherapy is the front-line treatment for schizophrenia, adjunctive treatments are needed since antipsychotic medications may not be sufficiently effective (Stahl, 1999) and noncompliance is a common problem (Ratakonda, Miller, Gorman, & Sharif, 1997). Bouchard, Vallieres, Roy, and Maziade (1996) reviewed 15 early studies on cognitive and behavioral interventions for delusions and hallucinations. Their summary indicated that CBT produced significant reductions in these symptoms. Since Bouchard et al.'s review there have been five randomized clinical trials published on CBT of schizophrenia. Rector and Beck (2001) reviewed this emerging literature and reported effect sizes. They concluded that CBT plus routine care (pharmacotherapy and case management) led to substantial pretreatment–posttreatment improvements in positive symptoms, negative symptoms, and total symptoms (grand mean ES=1.23). Patients receiving only routine care did much less well (grand mean ES=0.17). Alternative psychosocial interventions, such as supportive therapy or befriending, produced intermediate results (grand mean ES=0.60).

Gould, Mueser, Bolton, Mays, and Goff (2001), using results from seven controlled trials of CT for schizophrenia, performed a meta-analysis on the effects of CT on psychotic symptoms. The intervention consisted of targeted CBT strategies for modifying patients' distorted beliefs about delusions and hallucinations so as to decrease the negative consequences of these symptoms on their daily functioning. They found a large effect size for reduction in psychotic symptoms between pre- and post-treatment (grand mean ES=0.65). Of the seven studies, five reported significant differences in psychotic symptoms between the control and treatment conditions at post-treatment. The remaining two studies found non-significant decreases in psychotic symptoms.

With respect to follow-up findings, Rector and Beck (2001) found that across five studies, the follow-up effect size for change in frequency of and distress associated with delusions and hallucinations was large between pretest and follow-up (ES=1.48) and increased from treatment termination to follow-up. Similarly, large effect sizes were found for change in the magnitude of negative symptoms at 9-month follow-up across three studies (ES=0.88). Also, the small-to-moderate follow-up effect size for change in negative symptoms in the supportive therapy/regular care condition decreased between treatment termination and follow-up (from ES=0.47 to ES=0.22). In contrast, the effect size for the CBT condition, despite decreasing, remained large (ES=0.88) at follow-up. Similar to these findings, Gould et al. (2001) reported that in the four studies providing long-term follow-up findings, the gains made by patients who received CT were maintained and even increased in magnitude (ES=0.93).

2.8. Marital distress

Dunn and Schwebel (1995) conducted a meta-analytic review of marital therapy outcome research. They found a post-treatment ES of 0.71 for cognitive-behavioral marital therapy (CBMT) when compared to no treatment. This effect size was not significantly different from that of behavior marital therapy (BMT; ES=0.78) but was significantly less than that of interpersonally oriented marital therapy (IOMT; ES=1.37). The differences in effectiveness between CBMT, BMT, and IOMT narrowed at follow-up and were not statistically significant (ES's=0.54, 0.54, and 1.04, respectively). With respect to behavior change in treated couples, there were no significant differences in effect sizes for BMT, CBMT, and IOMT (ES=0.79, 0.54, and 0.87, respectively). Similarly, no significant differences were found at follow-up (ES=0.52, 0.75, 0.69), though it is interesting to note that the observed effect size for CBMT increased between termination and follow-up.

2.9. Anger

Beck and Fernandez (1998) reviewed 50 outcome studies on anger covering a total of 1640 subjects, many of whom were in programs for violent offenders. They found a mean weighted effect size of 0.70 for cognitive-behavioral therapy versus untreated controls. The average cognitive-behavioral therapy patients did better than 76% of untreated subjects on reduction of anger.

2.10. *Bulimia nervosa*

Whittal, Agras, and Gould (1999) examined 54 comparisons in their meta-analysis of psychosocial and pharmacological treatments for bulimia nervosa. They found that CBT was associated with substantial improvements in eating disordered behaviors and eating attitudes. Improvements with CBT were significantly larger than those for medication in terms of binge frequency, purge frequency, and eating attitudes (medication ES_u 's=0.66, 0.39, and 0.71, respectively). Parenthetically, there are very few controlled studies on cognitive-behavioral treatment for anorexia nervosa. While the cognitive-behavioral approach appears promising (see Wilson & Fairburn, 1998), a meta-analysis for this disorder is premature.

2.11. *Internalizing childhood disorders*

Grossman and Hughes (1992) reviewed and meta-analyzed 22 studies that used what they referred to as “self-control” interventions in treating children with depressive, anxious, or somatic disorders. These authors used the term “self-control therapy” as a synonym for cognitive-behavioral therapy and, in particular, for interventions “that target children’s distorted or deficient cognitive processing or teach coping responses” (p. 231). They found self-control therapy to be superior to no-treatment or placebo controls for depressive, anxious, and somatic disorders in children. The average primary effect size (0.74) was relatively large, and 77% of effects included were from studies measuring outcomes at follow-up (average duration=26 weeks).

2.12. *Sexual offending*

One of the more important outcome variables for sexual offenders is recidivism. Nagayama Hall (1995) conducted a meta-analysis of 12 recent treatment studies, 5 of which used a cognitive-behavioral approach. The subjects in these five studies included adults and adolescents charged with rape, attempted rape, or child molestation. Recidivism was defined somewhat differently across studies but all studies used official records to determine recidivism. The mean effect size of CBT compared to no treatment was 0.35, equal to that of hormonal therapy, and superior to behavioral therapy. Moreover, CBT showed lower refusal and drop out rates than hormonal treatments. While the CBT effect sizes were relatively small, given the impact on victims, it is arguable that any reduction in sexual offender recidivism is clinically meaningful. Of note, recidivism was measured from 1 year to as many as 28 years (mean=1.5 years) post-treatment, indicating long-lasting CT effects.

2.13. *Chronic pain*

In a recent meta-analysis, Morley, Eccleston, and Williams (1999) reviewed 25 trials in which the effectiveness of cognitive treatments for pain (not headache) was compared to wait list and alternative control conditions. They categorized dependent measures into domains including (a) pain experience, (b) mood, (c) cognitive coping and appraisal, (d) pain behavior/reduced activity, and (e) social role functioning. Cognitive treatments were significantly superior to waiting list on all domains except the expression of pain behavior, with a weighted grand mean ES of 0.46. Across domains, the effect sizes for cognitive treatments were generally larger than those for behavior therapy, although there were fewer trials with behavior therapy. There were too few studies using biofeedback to make statistical comparisons of effect sizes with CBT. However, the effects sizes reported by Morley et al. for biofeedback, in the area of 0.70, suggest that this is a promising intervention for chronic pain. When CBT was compared to a heterogeneous collection of alternative treatments (not including behavior therapy or biofeedback), it was found to be superior in reducing pain experience, increasing positive cognitive coping and appraisal, and reducing behavioral expression of pain. No data are reported concerning long-term effectiveness of treatment.

3. Summary and conclusions

In this review of 16 recent meta-analyses we sought to answer a multifaceted question: How effective is CBT, for which disorders, and compared to what, and how lasting are these effects? Collectively, the findings detailed in

this review suggest that CT is highly effective for adult unipolar depression, adolescent unipolar depression, generalized anxiety disorder, panic disorder with or without agoraphobia, social phobia, PTSD, and childhood depressive and anxiety disorders. The comparison-weighted grand mean effect size for these disorders when compared to no-treatment, wait list, or placebo controls is 0.95 ($SD=0.08$). CBT is associated with large improvements in symptoms for bulimia nervosa, and the associated effect sizes ($M=1.27$, $SD=0.11$) are significantly larger than those that have been found for pharmacotherapy. CBT also has shown promising results as an adjunct to pharmacotherapy in the treatment of schizophrenia. The average uncontrolled effect size of 1.23 for CBT compares favorably with an ES of 0.17 for schizophrenic patients receiving only routine care. Moderate effect sizes ($M=0.62$, $SD=0.11$) were obtained when CT was compared to controls for marital distress, anger, childhood somatic disorders, and several chronic pain variables (i.e., pain expression behavior, activity level, social role functioning and cognitive coping and appraisal). CT was somewhat superior to antidepressants in the treatment of adult unipolar depression ($ES=0.38$). CT was equally effective to behavior therapy in the treatment of adult depression ($ES=0.05$) as well as obsessive-compulsive disorder ($ES=0.19$). Trauma-focused CBT and EMDR were equally effective for PTSD. The efficacy of CT for sexual offending is relatively low ($ES=0.35$). However, along with hormonal treatments, it is the most effective treatment for reducing recidivism in this population. Finally, CT was found to be superior to supportive/nondirective therapy in the few occurrences when these two treatments were directly compared. This includes two comparisons for adolescent depression ($ES=0.84$) and two comparisons for generalized anxiety disorder ($ES=0.71$).

We also aimed to provide answers regarding the degree to which the effects of CT persist following the termination of treatment. The meta-analyses reviewed strongly suggest that across many disorders the effects of CT are maintained for substantial periods beyond the cessation of treatment. More specifically, significant evidence for long-term effectiveness was found for depression, generalized anxiety, panic, social phobia, OCD, sexual offending, schizophrenia, and childhood internalizing disorders. In the cases of depression and panic, there appears to be robust and convergent meta-analytic evidence that CT produces vastly superior long-term persistence of effects, with relapse rates half those of pharmacotherapy. In addition, CT appears to show greater long-term effects in the treatment of generalized anxiety disorder as compared to applied relaxation.

This review of existing meta-analyses adds to our knowledge of the effects of CT for a variety of disorders. However, future meta-analyses should (1) continue to evaluate the effects of CBT as applied to an even wider, more specific spectrum of disorders, (2) continue to evaluate the long-term effects of CBT, and (3) continue to evaluate the effects of CBT as compared to alternative treatments. For example, with respect to the former, Lam et al. (2003), in a randomized controlled trial of CBT for relapse prevention in Bipolar Disorder, found that patients treated with CBT experienced significantly better outcomes (e.g., fewer and shorter bipolar episodes, fewer hospital admissions, less variability in manic symptoms, etc.) at 1 year follow-up. However, because there are currently no meta-analytic findings on CBT for Bipolar Disorder, the application of meta-analytic procedures will make an important contribution as similar studies are completed.

Similarly, for many disorders the current review offers little to no evidence bearing on the question of the persistence of CBT effects relative to other psychological and pharmacological interventions. The degree to which CBT is or is not superior to alternative treatments in terms of follow-up findings has only been partially addressed by the meta-analyses reviewed here. For example, limited findings were available on the long-term effects of CBT for anger, bulimia, anorexia nervosa, and chronic pain. The current evidence, where available, does suggest a general advantage in favor of CBT. For disorders lacking such findings, we incorporated recent reviews of CBT. For example, the review by DeRubeis and Crits-Cristoph (1998) provided additional support for the long-term efficacy of CBT for depression, GAD, and social phobia.

Clearly, more research is needed comparing CBT with other forms of therapy. Indeed, the effect sizes we reported comparing CBT to other active treatments for adolescent depression, generalized anxiety disorder, panic disorder, and obsessive-compulsive disorder must be considered preliminary due to the low number of studies in the corresponding meta-analyses. We included these meta-analyses in this review since the effect sizes reported provide a starting point for evaluating the stability of results across samples, designs, and analyses. Such information also informs power analyses for future primary studies (Wilkinson, 1999).

Future meta-analyses will continue to face the challenges associated with long-term findings such as variable follow-up assessment time-points across studies and the handling of data for patients who drop out of studies during the follow-up period (Dobson, 1989). Despite these obstacles, the examination of long-term effects has critical

implications for the viability of treatments, treatment recommendations and decision making, and the determination of cost-benefit ratios and economic efficiency.

Meta-analysis has gained increasing recognition as a valid procedure for evaluating the efficacy of a treatment (Cooper & Hedges, 1994; Wilkinson, 1999). However, the limitations of this approach should be kept in mind when drawing conclusions from this review. The practice of reporting multiple effect sizes for individual studies is problematic since these are typically not independent. This makes it likely that some of the studies are given more weight than others. Also, simply aggregating effect sizes may mask important variability among studies in methodological rigor, sample characteristics, and measurement instruments and procedures. In some cases, such factors can be quantitatively assessed and incorporated into a meta-analysis as moderator variables. Where possible, we reported results on important moderator variables that have been studied meta-analytically (e.g., researcher allegiance in depression studies). However, inclusion of moderator variables was not generally the case in the meta-analyses we reviewed. Another limitation of averaging effect sizes is that striking successes and failures may be overlooked and these are important in interpreting outcome findings. In this respect, it is important to note that the conclusions from this meta-analysis review are consistent with recent reviews of high-quality primary studies (DeRubeis & Crits-Christoph, 1998; Nathan & Gorman, 1998).

There has been a trend since the 1970s to apply CBT to an increasingly wider spectrum of disorders (Beck, 1997). Substance abuse, bipolar disorder, personality disorders, and anorexia nervosa are among the disorders receiving recent empirical attention. As additional controlled outcome studies accumulate, future meta-analyses will be needed to examine the strength of CT in the treatment of these and other disorders.

References

- Abramowitz, J. S. (1997). Effectiveness of psychological and pharmacological treatments for obsessive-compulsive disorder: A quantitative review. *Journal of Consulting and Clinical Psychology, 65*, 44–52.
- Antonuccio, D. O., Thomas, M., & Danton, W. G. (1997). A cost-effectiveness analysis of cognitive behavior therapy and fluoxetine (prozac) in the treatment of depression. *Behavior Therapy, 28*, 187–210.
- Beck, A. T. (1997). The past and future of cognitive therapy. *Journal of Psychotherapy Practice and Research, 6*, 276–284.
- Beck, J. S. (1995). *Cognitive therapy: Basics and beyond*. New York: Guilford Press.
- Beck, R., & Fernandez, E. (1998). Cognitive-behavioral therapy in the treatment of anger: A meta-analysis. *Cognitive Therapy and Research, 22*, 63–74.
- Beutler, L. E. (2002). The dodo bird is extinct. *Clinical Psychology: Science & Practice, 9*, 30–34.
- Blackburn, I. M., Bishop, S., Glen, A. I., Whalley, L. J., & Christie, J. E. (1981). The efficacy of cognitive therapy in depression: A treatment trial using cognitive therapy and pharmacotherapy, each alone and in combination. *British Journal of Psychiatry, 139*, 181–189.
- Blatt, S. J., Zuroff, D. C., Bondi, C. M., & Sanislow, C. A. (2000). Short- and long-term effects of medication and psychotherapy in the brief treatment of depression: Further analyses of data from the NIMH TDCRP. *Psychotherapy Research, 10*(2), 215–234.
- Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 61*, 611–619.
- Bouchard, S., Vallieres, A., Roy, M. A., & Maziade, M. (1996). Cognitive restructuring in the treatment of psychotic symptoms in schizophrenia: A critical analysis. *Behavior Therapy, 27*, 257–277.
- Chambless, D. L. (2002). Beware the dodo bird: The dangers of overgeneralization. *Clinical Psychology: Science & Practice, 9*, 13–16.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H. M., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.
- DeRubeis, R. J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology, 66*, 37–52.
- DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., et al. (2005). Cognitive therapy vs. medications in the treatment of moderate to severe depression. *Archives of General Psychiatry, 62*, 409–416.
- Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 57*, 414–419.
- Dobson, K. S. (Ed.). (2001). *Handbook of cognitive-behavioral therapies*, (2nd ed.). New York: Guilford Press.
- Dobson, K. S., & Dozois, D. J. (2001). Historical and philosophical bases of the cognitive-behavioral therapies. In K. S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (2nd ed.). New York: Guilford Press.
- Dunn, R. L., & Schwebel, A. I. (1995). Meta-analytic review of marital therapy outcome research. *Journal of Family Psychology, 9*, 58–68.
- Durham, R. C., Chambers, J. A., MacDonald, R. R., Power, K. G., & Major, K. (2003). Does cognitive-behavioural therapy influence the long-term outcome of generalized anxiety disorder? An 8–14 year follow-up of two clinical trials. *Psychological Medicine, 33*, 499–509.
- Elkin, I., Shea, M. T., Watkins, S. D., Imber, S. M., Sotsky, S. M., Collins, D. R., et al. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971–982.

- Feske, U., & Chambless, D. L. (1995). Cognitive behavioral versus exposure only treatment for social phobia: A meta-analysis. *Behavior Therapy*, 26, 695–720.
- Gaffan, E. A., Tsaousis, I., & Kemp-Wheeler, S. M. (1995). Research allegiance and meta-analysis: The case of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 63, 966–980.
- Glass, G. V., McGraw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Gloaguen, V., Cottraux, J., Cucherat, M., & Blackburn, I. (1998). A meta-analysis of the effects of cognitive therapy in depressed patients. *Journal of Affective Disorders*, 49, 59–72.
- Gould, R. A., Buckminster, S., Pollack, M. H., Otto, M. W., & Yap, L. (1997). Cognitive-behavioral and pharmacological treatment for social phobia: A meta-analysis. *Clinical Psychology: Science & Practice*, 4, 291–306.
- Gould, R. A., Mueser, K. T., Bolton, E., Mays, V., & Goff, D. (2001). Cognitive therapy for psychosis in schizophrenia: An effect size analysis. *Schizophrenia Research*, 48, 335–342.
- Gould, R. A., Otto, M. W., & Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review*, 15(8), 819–844.
- Gould, R. A., Otto, M. W., Pollack, M. H., & Yap, L. (1997). Cognitive behavioral and pharmacological treatment of generalized anxiety disorder: A preliminary meta-analysis. *Behavior Therapy*, 28, 285–305.
- Grossman, P. B., & Hughes, J. N. (1992). Self-control interventions with internalizing disorders: A review and analysis. *School Psychology Review*, 21(2), 229–245.
- Hollon, S. D., & Beck, A. T. (1994). Cognitive and cognitive-behavioral therapies. In A. E. Bergin, & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change*, (4th ed.) (pp. 428–466).
- Hollon, S. D., DeRubeis, R. J., Shelton, R. C., Amsterdam, J. D., Salomon, R. M., & O'Reardon, J. P., et al. (2005). Prevention of relapse following cognitive therapy vs. medications in moderate to severe depression. *Archives of General Psychiatry*, 62, 417–422.
- Hollon, S. D., Thase, M. E., & Markowitz, J. C. (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest*, 3, 39–77.
- Howard, K. I., Krause, M. S., Saunders, S. M., & Kopta, S. M. (1997). Trials and tribulations in the meta-analysis of treatment differences: Comment on Wampold et al. (1997). *Psychological Bulletin*, 122, 221–225.
- Lam, D. H., Watkins, E. R., Harward, P., Bright, J., Wright, K., & Kerr, N., et al. (2003). A randomized controlled study of cognitive therapy for relapse prevention for bipolar affective disorder: Outcome of the First Year. *Archives of General Psychiatry*, 60(2), 145–152.
- Luborsky, L., Rosenthal, R., Diguier, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., et al. (2002). The dodo bird verdict is alive and well—mostly. *Clinical Psychology: Science & Practice*, 9, 2–12.
- Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science & Practice*, 9, 21–25.
- Morley, S., Eccleston, C., & Williams, A. (1999). Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. *Pain*, 80, 1–13.
- Nagayama Hall, G. C. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology*, 63(5), 802–809.
- Nathan, P. E., & Gorman, J. M. (1998). *A guide to treatments that work*. Oxford: Oxford University Press.
- Oei, T. P. S., Llamas, M., & Devilly, G. J. (1999). The efficacy and cognitive processes of cognitive behaviour therapy in the treatment of panic disorder with agoraphobia. *Behavioural and Cognitive Psychotherapy*, 27, 63–88.
- Olfson, M., & Klerman, G. L. (1993). Trends in the prescription of antidepressants by office-based psychiatrists. *American Journal of Psychiatry*, 150, 571–577.
- Parker, G., Roy, K., & Eyers, K. (2003). Cognitive behavior therapy for depression? Choose horses for courses. *American Journal of Psychiatry*, 160, 825–834.
- Paykel, E. S., Scott, J., Teasdale, J. D., Johnson, A. L., Garland, A., Moore, R., et al. (1999). Prevention of relapse in residual depression by cognitive therapy: A controlled trial. *Archives of General Psychiatry*, 56, 829–835.
- Ratakonda, S., Miller, C. E., Gorman, J. M., & Sharif, Z. A. (1997). Efficacy of a 12-week trial of olanzapine in treatment of refractory schizophrenia or schizoaffective disorder. *Schizophrenia Bulletin*, 29, 151.
- Rector, N. A., & Beck, A. T. (2001). Cognitive-behavioral therapy for schizophrenia: An empirical review. *Journal of Nervous and Mental Disease*, 189, 278–287.
- Rector, N. A., Richter, M. A., Denisoff, E., Crawford, C., Szacun-Shimizu, K., Bourdeau, D. (2005). *Cognitive and behavioral treatments for medication-refractory OCD*.
- Reinecke, M. A., Ryan, N. E., & DuBois, D. L. (1998). Cognitive-behavioral therapy of depression and depressive symptoms during adolescence: A review and meta-analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37(1), 26–34.
- Rounsaville, B. J., & Carroll, K. M. (2002). Commentary on dodo bird revisited: Why aren't we dodos yet? *Clinical Psychology: Science & Practice*, 9(1), 17–20.
- Rush, A. J., Beck, A. T., Kovacs, M., & Hollon, S. D. (1977). Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. *Cognitive Therapy and Research*, 1(1), 7–37.
- Salkovskis, P. M. (Ed.). (1996). *Frontiers of cognitive therapy*. New York: Guilford Press.
- Shapiro, D. A., Barkham, M., Rees, A., Hardy, G. A., Reynolds, S., & Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 522–534.
- Simons, A. D., Levine, J. L., Lustman, P. J., & Murphy, G. E. (1984). Patient attrition in a comparative outcome study of depression: A follow-up report. *Journal of Affective Disorders*, 6, 163–173.

- Stahl, S. M. (1999). Selecting atypical antipsychotics by combining clinical experience with guidelines from clinical trials. *Journal of Clinical Psychiatry*, *60*, 31–41.
- Thase, E. M., Greenhouse, J. B., Frank, E., Reynolds, C., Palkonis, P. A., Hurley, K., et al. (1997). Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations. *Archives of General Psychiatry*, *54*, 1009–1015.
- Treatment for Adolescents With Depression Study (TADS) Team, US. (2004). Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for adolescents with depression study (TADS) randomized controlled trial. *Journal of the American Medical Association*, *292*(7), 807–820.
- Van Balkom, A. J. L. M., van Oppen, P., Vermeulen, A. W. A., van Dyck, R., Nauta, M. C. E., & Vorst, H. C. M. (1994). A meta-analysis on the treatment of obsessive-compulsive disorder: A comparison of antidepressants, behavior, and cognitive therapy. *Clinical Psychology Review*, *14*, 359–381.
- Wampold, B. E., Minami, T., Baskin, T. W., & Tierney, S. C. (2002). A meta-(re) analysis of the effects of cognitive therapy versus ‘other therapies’ for depression. *Journal of Affective Disorders*, *68*, 159–165.
- Whittal, M. L., Agras, W. S., & Gould, R. A. (1999). Bulimia nervosa: A meta-analysis of psychological and pharmacological treatments. *Behavior Therapy*, *30*, 117–135.
- Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.
- Wilson, G. T., & Fairburn, C. G. (1998). Treatments for eating disorders. In P. E. Nathan, & J. M. Gorman (Eds.), *A guide to treatments that work*. Oxford: Oxford University Press.