CORRECTED AUGUST 22, 2008; SEE LAST PAGE

Journal of Consulting and Clinical Psychology
2007, Vol. 75, No. 4, 513–522

0022-006X/07/$12.00 DOI: 10.1037/0022-006X.75.4.513

Specificity of Treatment Effects: Cognitive Therapy and Relaxation for Generalized Anxiety and Panic Disorders

Jedidiah Siev and Dianne L. Chambless
University of Pennsylvania

The aim of this study was to address claims that among bona fide treatments no one is more efficacious than another by comparing the relative efficacy of cognitive therapy (CT) and relaxation therapy (RT) in the treatment of generalized anxiety disorder (GAD) and panic disorder without agoraphobia (PD). Two fixed-effects meta-analyses were conducted, for GAD and PD separately, to review the treatment outcome literature directly comparing CT with RT in the treatment of those disorders. For GAD, CT and RT were equivalent. For PD, CT, which included interoceptive exposure, outperformed RT on all panic-related measures, as well as on indices of clinically significant change. There is ample evidence that both CT and RT qualify as bona fide treatments for GAD and PD, for which they are efficacious and intended to be so. Therefore, the finding that CT and RT do not differ in the treatment of GAD, but do for PD, is evidence for the specificity of treatment to disorder, even for 2 treatments within a CBT class, and 2 disorders within an anxiety class.

Keywords: generalized anxiety disorder, panic disorder, cognitive therapy, relaxation therapy, treatment specificity

There has been extensive debate and controversy in recent years regarding the utility of examining active ingredients versus common factors in psychotherapy. Wampold and colleagues (e.g., Ahn & Wampold, 2001; Wampold et al., 1997) and Luborsky and colleagues (e.g., Luborsky et al., 2002) have claimed that all forms of psychotherapy are equivalent; what matters are the quality of the therapeutic alliance and other factors common to all treatments such as the presentation of a treatment rationale. Their assertions are based on meta-analyses that combine treatments of all types and disorders of all types, largely for adult patients. Researchers have challenged their conclusions on the grounds that the fact that all treatments for all disorders when combined do not differ from each other does not imply that a particular treatment for a particular disorder is not superior (Chambless, 2002). Proponents of a common factors approach to psychotherapy limit their contentions to *bona fide treatments*, defined as "those that were delivered by trained therapists and were based on psychological principles, were offered to the psychotherapy community as viable treatments (e.g., through professional books or manuals), or contained specified components" (Wampold et al., 1997, p. 205). This reasoning, however, presupposes that a treatment can be, in and of itself, bona fide, with no clinical referent: for the treatment of what?

Ahn and Wampold (2001) noted that "a familiar criticism of meta-analysis" is that "aggregating across diverse studies yields spurious conclusions" (p. 254). Certainly in conducting a meta-analysis, the researcher paints with broad strokes. However, if the findings are to be clinically useful, the researcher must consider meaningful subsets in the data so as not to obscure potentially important differences among treatments. It is our argument that significant differences among treatments may exist for specific disorders. For example, there are at least four studies in which exposure and response prevention for obsessive–compulsive disorder (OCD) has proven more efficacious than relaxation therapy (RT; see Chambless & Ollendick, 2001).

The primary goal of this article is to address this specific aspect of the common factors approach by evaluating two efficacious, active treatments across studies, but within particular domains of psychopathology. A second goal, which forms the framework for this investigation, is to review the treatment outcome literature comparing cognitive therapy (CT) and RT for generalized anxiety disorder (GAD) and panic disorder without agoraphobia (PD) to determine the relative efficacy of each treatment for each disorder. GAD and PD are related, and, until the introduction of the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.; American Psychiatric Association, 1980), were not diagnostically distinct. In fact, even subsequent to the formalization of that distinction, some studies evaluating cognitive–behavioral therapy (CBT) and RT in treating anxiety have used mixed GAD–PD samples (e.g., Borkovec & Mathews, 1988; Öst, 1988). By analyzing particular treatments in particular domains, it is possible accurately to compare two specific active ingredients in psychotherapy, and by doing so within the context of two related anxiety disorders, one executes a conservative comparison. Although the absence of significant differences between treatments would not necessarily imply that common factors account for all effects, the presence of differences would support the notion that treatment effects are influenced by specific therapy techniques.

Jedidiah Siev and Dianne L. Chambless, Department of Psychology, University of Pennsylvania.

We wish to thank Arnoud Arntz and Lars-Göran Öst for generously sharing data to facilitate this meta-analysis. We are grateful, as well, to Arnoud Arntz for suggesting this investigation.

Correspondence concerning this article should be addressed to Jedidiah Siev, Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104-6241. E-mail: jsiev@psych.upenn.edu

The reasons GAD and PD were selected from the anxiety disorders as domains of investigation are as follows. Similar studies of panic disorder with agoraphobia (PDA) have all included in vivo exposure as an element of CBT, and for anxiety disorders, in vivo exposure typically renders supplemental CT unnecessary (e.g., Öst, Thulin, & Ramnerö, 2004). In contrast, most PD studies have included interoceptive exposure in the CT condition, but not in vivo exposure to phobic situations. CBT as administered in studies of PDA was therefore judged to be qualitatively different from that used in studies of GAD and PD. Studies of OCD were also not included, not only because the CBT condition in CBT versus RT studies for OCD consisted almost entirely of exposure and response prevention, but also because RT was intended as a control condition in those studies and might not satisfy the criteria of a bona fide treatment. GAD and PD are the only other anxiety disorders for which multiple randomized clinical trials have been conducted that compare CT and RT directly.

CT and RT for GAD and PD

GAD and PD are common and costly. The National Comorbidity Survey found lifetime prevalence rates of approximately 5% and 3.5% for GAD and PD, respectively (Kessler et al., 1994; Wittchen, Zhao, Kessler, & Eaton, 1994). Furthermore, GAD is often chronic, resistant to change, and characterized by early onset (Sanderson & Wetzler, 1991; Zuellig & Newman, 1996, as cited by Borkovec, Newman, Pincus, & Lytle, 2002). The substantial individual health and social costs associated with PD rival those connected with major depressive disorder, and PD patients utilize medical services at exceptionally high rates (Katon, 1996; Markowitz, Weissman, Ouellette, Lish, & Klerman, 1989). Although it is perhaps the basic anxiety disorder (Brown, Barlow, & Liebowitz, 1994), GAD has proven particularly difficult to treat, with clinical trials producing clinically significant improvement in only about 50% of participants (Borkovec & Newman, 1998; Borkovec & Whisman, 1996). Treatments for PD, in contrast, have produced clinically significant improvement in more than 80% of research participants (Clark, 1996; Öst & Westling, 1995).

Among other psychotherapy approaches, CT and RT have received considerable empirical support in the treatment of anxiety disorders, and GAD and PD in particular (e.g., Gould, Otto, & Pollack, 1995; Gould, Otto, Pollack, & Yap, 1997; Mitte, 2005a, 2005b). Variations of the two are often combined in CBT packages, but they are based, in principle, on different theoretical approaches, and each one individually appears efficacious in clinical trials. According to the cognitive model, anxiety is maintained by an individual's misperception of danger and catastrophic misinterpretations of generally benign stimuli, sometimes internal and other times external. For example, a stimulus might precipitate a physiological or cognitive reaction from an individual, which is followed by an exaggerated perception of danger or threat, which strengthens the reaction, and so forth, creating a self-perpetuating cycle of progressively intensified anxiety. Adherents to the cognitive model propose that these thoughts are accessible to conscious consideration and intentional reevaluation, staples of CT (e.g., A. T. Beck, Emery, & Greenberg, 1985; Clark, 1986). The applications of the cognitive model to GAD and PD are discussed more specifically later, in terms of the treatment protocols used by the authors of studies analyzed herein.

RT is a coping technique with a behavioral treatment component, whereby individuals learn to relax in the presence or anticipation of feared stimuli or general anxiety (e.g., Öst, 1987, 1988). Proponents of RT subscribe to a model of anxiety similar to the cognitive model, but the focus in treatment is entirely different. The goal is to abort the anxiety cycle by decreasing the intensity of the physiological reaction and to avoid catastrophic thoughts without addressing the cognitions directly. Although there is limited empirical evidence for particular mechanisms of change, it is believed that RT works by (a) reducing general tension and anxiety and, consequently, the likelihood that any particular stressor will trigger panic; (b) increasing awareness about how anxiety works, thereby demystifying it and diminishing its impact; and (c) enhancing perceived self-efficacy, so that the individual feels equipped to cope with an anxiety reaction (Öst, 1987, 1992). This approach is consistent both with the cognitive model and with behavioral models of anxiety in which anxious processes are viewed as primarily automatic and out of consciousness, and successful treatment as operating at that level (e.g., Marks, 1987; Ohman & Soares, 1994).

There are two primary forms of RT investigated in the GAD and PD treatment outcome literatures during the past 3 decades: *progressive relaxation* (PR; e.g., Bernstein & Borkovec, 1973) and *applied relaxation* (AR; Öst, 1987). During the former, which also composes the first stage of AR training, the individual learns to relax by tensing and then relaxing various muscle groups. After one is able to achieve a state of relaxation in that way, one learns to relax without first physically tensing muscles, to relax in response to a self-generated cue, and finally to relax those muscles not involved in a particular activity even as one engages in that activity (*differential relaxation*). Öst (1987) added *application training* to PR, in which individuals practice applying the relaxation in vivo by approaching increasingly feared situations and using the relaxation techniques to manage the evoked anxiety. Although AR, by definition, involves exposure, the goal is not to extinguish the anxiety through a process of habituation, but rather to practice applying the skills in vivo. In fact, exposures are often not sufficiently long to allow for habituation.

Although a number of previous meta-analyses have evaluated the efficacy of CBT for GAD and PD (e.g., Mitte, 2005a, 2005b), they did not focus specifically on the relative efficacy of CT and RT, two known active treatments. Not only is there considerable evidence that these therapies work, both in combination with each other and separately, but a number of studies have now been published that directly compare the two active treatments for GAD and PD (see references marked with an asterisk). It is important, scientifically and clinically, that the results of these trials be synthesized to describe the state of the literature with respect to these treatments.

Of specific interest as outcome measures are those related to core features of the psychopathology of each disorder: generalized anxiety and anxiety-related cognitions in the case of GAD, and panic, fear of anxiety, and panic-related cognitions in the case of PD. Although other symptoms such as depression are certainly of interest, it is clinically most important to determine whether treatments differ in impact on the patients' core presenting problems.

Meta-Analytic Procedure

Standardized Mean Difference

The standardized mean difference effect size was calculated from posttreatment data to evaluate between-groups differences based on their respective mean scores. This effect size was adjusted for a slight upward bias, yielding Hedges' *g* (Hedges, 1981). As reported in the present article, a positive sign denotes that an effect size favors CT over RT, whereas a negative sign indicates a relative advantage for RT over CT.

In addition, each effect size in a given analysis was weighted to account for its relative precision, in large part a function of sample size. Hedges and Olkin (1985) suggested basing the weights on the standard errors of the effect sizes, rather than simply the sample size. Specifically, the inverse variance weight was used, which is the reciprocal of the squared standard error:

$$w = \frac{1}{SE^2}.$$

Fixed-effects meta-analyses proceeded as follows. First, the effect size, standard error, and inverse variance weight were calculated for each construct measured within each study. Second, the weighted mean effect size was computed for a given domain across all studies, according to the formula

$$\overline{ES} = \frac{\sum (w_i \times ES_i')}{\sum w_i},$$

as was the standard error of the weighted mean effect size,

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum w_i}}.$$

Often, individual studies report multiple dependent measures of a single construct. Including more than one measure per study would violate assumptions of independence, inflate the sample size, and distort standard error estimates. Hence, a single effect size for each domain within each study was calculated by taking the arithmetic mean of the effect sizes for all included measures in that domain. The standard error was computed by treating the mean effect size as a single effect size from which to calculate a standard error; it was not the mean of the standard errors. This method has been demonstrated to produce appropriate estimates when homogeneity of effect sizes is assumed, as is the case in fixed-effects analyses (Marin-Martinez & Sanchez-Meca, 1999).

Odds Ratio

Dichotomous variables were analyzed using the odds ratio effect size, which denotes the ratio of odds for a particular outcome for two groups, where the odds of an event is the probability of an outcome divided by 1 minus that probability. If any cell frequency was equal to 0, 0.5 was added to all cells. An odds ratio of 1 indicates identical odds for the two comparison groups. Because the distribution is skewed, analyses were performed on the natural log of the odds ratio. The inverse variance weight and weighted mean effect size were calculated by the same formulas as when computing the standardized mean difference. For ease of interpretation, summary statistics were then transformed back to odds ratios by taking antilogarithms.

Proportion

Weighted average proportions were calculated to give the percentage of each group that met certain criteria. Because proportions are constrained to values between 0 and 1, using simple proportions can distort the results, especially when the mean proportion approaches 0 or 1. Therefore, proportions were converted into logits as the effect sizes, analyses were conducted on the logits, and the results were converted back to proportions. Again, the same formulas as described earlier were applied to compute the weighted effect size.

Homogeneity Analysis

Fixed-effects models are predicated on the assumption that the variance in effect sizes among the studies can be attributed to sampling error. That is, a homogeneous distribution describes a set of effect sizes that are dispersed around their mean as would be expected due only to subject-level sampling error. A heterogeneous distribution suggests that there is additional variability, either random or systematic, that usually warrants modeling, given sufficient power. It is possible statistically to test whether there is evidence of heterogeneity. This test is based on the *Q* statistic, which has a chi-square distribution (Hedges & Olkin, 1985). A significant *Q* statistic suggests heterogeneity.

Because relatively few studies were included in this meta-analysis, fixed-effects analyses were preferred to avoid Type II error. However, *Q* has poor power to detect heterogeneity, so a conservative level of $\alpha = .10$ was set to denote significant heterogeneity. In the case of significant homogeneity tests, random-effects analyses were planned to model the between-studies variability; for nonsignificant homogeneity tests, fixed-effects analyses were used. Recently, Higgins and Thompson (2002) proposed an index, I^2, which can be interpreted as the percentage of variability due to heterogeneity; it therefore contains information about the degree of heterogeneity, not just significance level (see also Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). Higgins and Thompson offered the tentative heuristic that "mild heterogeneity might account for less than 30 per cent of the variability in point estimates, and notable heterogeneity substantially more than 50 per cent" (p. 1553). Similarly, Hudeo-Medina et al. referred to a classification of low, medium, and high heterogeneity corresponding to I^2 values of 25, 50, and 75.

GAD

CBT for GAD typically comprises some or all of the following elements: (a) self-monitoring; (b) cognitive restructuring, including evaluating and reconsidering interpretive and predictive cognitions; (c) relaxation training; and (d) rehearsal of coping skills (Borkovec & Ruscio, 2001). Because GAD is not strongly characterized by avoidance of specific external environments or situations, CBT for GAD makes use of imagery rehearsal (e.g., stress inoculation) more than the in vivo exposure that would be fundamental to CBT for other anxiety disorders (Borkovec & Ruscio, 2001). As previously mentioned, the efficacy of CBT packages for

GAD is well documented, if less impressive than for other disorders (e.g., Roemer & Orsillo, 2002).

Studies were located via a search of PsycINFO using the keywords *cognitive therapy* or *cognitive–behavioral therapy* and *generalized anxiety disorder.* In addition, the major journals publishing CBT outcome studies were checked by hand from 1992 to 2005: *Behavior Therapy*, *Behaviour Research and Therapy*, *Cognitive Therapy and Research*, *Journal of Anxiety Disorders*, and *Journal of Consulting and Clinical Psychology*. Last, the text and reference sections of relevant studies, articles, and chapters were scanned. Studies were included if they compared therapist-administered, individual CT and RT; participants met diagnostic criteria for GAD and were randomized to treatment condition; and the two treatment protocols were judged not to contain overlapping ingredients presumed to be active. For example, whereas structural features such as format and length were expected to be similar, four studies were excluded primarily because the CT condition included relaxation training or similar behavioral techniques. Including only studies that compared CT and RT directly, although limiting in number, is advantageous because it minimizes apparent treatment differences that are artifacts of cross-study particularities (see Weisz, Jensen-Doss, & Hawley, 2006). Five studies were identified that directly compared CT and RT (Arntz, 2003; Barlow, Rapee, & Brown, 1992; Borkovec et al., 2002; Butler, Fennell, Robson, & Gelder, 1991; Öst & Breitholtz, 2000); a sixth was in the stage of data collection and was unavailable for this meta-analysis (Dugas & Koerner, 2005).

By and large, the five studies were methodologically sound. With the exception of a subset of the participants in Arntz (2003), all participants were screened on the basis of validated structured interviews, and reanalysis of that study excluding other participants did not change the results. Also with the exception of Arntz (2003), who collected only self-report data, participants in all studies were assessed by independent raters blind to treatment condition. All but Öst and Breitholtz (2000) reported that therapists received regular supervision to enhance quality and adherence to treatment protocol, and three formally assessed integrity with excellent results (Barlow et al., 1992; Borkovec et al., 2002; Butler et al., 1991). Perceived treatment credibility or client expectations were assessed by all studies except Arntz (2003); they appeared rather high, and in no case differed by treatment group. The mean age of study participants was consistent across studies (approximately 33–40 years old), and in all studies therapists were crossed with treatment condition.

Overall, there was little evidence across studies for the superiority of either treatment over the other; only Butler et al. (1991) found that CT outperformed RT. One of the methodological limitations common to the studies, however, is the relatively small sample size and consequent low power to detect group differences. To address this concern and synthesize the various findings quantitatively, data were meta-analyzed using completer samples because data for intention-to-treat samples were not consistently available. Follow-up data were also not consistently available. Measures were categorized in three domains: primary domains of anxiety and anxiety-related cognitions, and the secondary domain of depression, which often occurs concurrently with anxiety. Effect sizes were calculated for between-groups differences at posttreatment. Particular measures that composed each domain for each study are given in Table 1. In addition, effect sizes were calculated to assess between-groups differences in rates of clinically significant change and drop-outs. As referred to herein, *clinically significant change* includes both measures of change and *high end-state functioning*, which refers to the quality of functioning at posttreatment.

Results

Between-groups comparisons. Between-groups effect sizes were calculated from uncontrolled posttreatment means and standard deviations. Unpublished data were obtained (A. Arntz, personal communication, January 27, 2006) to supplement those

Table 1
Measures Included in Generalized Anxiety Disorder Meta-Analysis

Study	Anxiety	Cognitive	Depression
Arntz (2003)	Fear of Fear (van den Hout et al., 1987), SCL-90 Anxiety, STAI-T		Bouman
Barlow et al. (1992)	ADIS-R Severity, CSAQ-Somatic, HARS, STAI-T, average daily anxiety, intense anxiety episodes per week	CSAQ-Cognitive, percentage of day spent worrying	BDI, HDRS, daily depression
Borkovec et al. (2002)	ADIS-R Severity, HARS, STAI-T, diary severity	PSWQ	BDI, HDRS
Butler et al. (1991)	BAI, HARS, Leeds Anxiety, STAI-T, Watson & Marks (1971), self-rated anxiety	SPQ, Interpretations, Individual Thoughts	BDI, Leeds Depression, Watson & Marks (1971), self-rated depression
Öst & Breitholtz (2000)	ADIS-R Severity, BAI, CSAQ-Somatic, HARS, STAI-T, self-rated severity	PSWQ, CSAQ-Cognitive	BDI, HDRS

Note. SCL-90 Anxiety = Symptom Checklist-90 Anxiety—Dutch version (Arrindell & Ettema, 1981); STAI-T = State-Trait Anxiety Inventory—Trait scale (Spielberger et al., 1970); Bouman = Bouman Depression Inventory (Bouman, 1987); ADIS-R = Anxiety Disorders Interview Schedule— Revised (DiNardo & Barlow, 1988); CSAQ = Cognitive Somatic Anxiety Questionnaire (Schwartz et al., 1978); HARS = Hamilton Anxiety Scale (Hamilton, 1959); BDI = Beck Depression Inventory (A. T. Beck et al., 1961); HDRS = Hamilton Depression Rating Scale (Hamilton, 1960); PSWQ = Penn State Worry Questionnaire (Meyer et al., 1990); BAI = Beck Anxiety Inventory (A. T. Beck et al., 1988); Leeds = Leeds Scales (Snaith et al., 1976); SPQ = Subjective Probabilities Questionnaire (Butler & Mathews, 1983).

printed in Arntz (2003). See Table 2 for between-groups effect sizes.

There was little difference between the effects of CT and RT on anxiety, anxiety-related cognitions, and depression, for which the weighted mean effect sizes were all small and nonsignificant: for anxiety, $g = 0.11$, $p = .32$, 95% confidence interval (CI) = −0.19, 0.40; for anxiety-related cognitions, $g = 0.13$, $p = .47$, CI = −0.21, 0.46; for depression, $g = 0.05$, $p = .73$, CI = −0.25, 0.35.

Homogeneity analyses were all nonsignificant: for anxiety, $Q(4) = 3.64$, $p = .46$, $I^2 = 0$; for anxiety-related cognitions, $Q(3) = 4.93$, $p = .18$, $I^2 = 39$; and for depression, $Q(4) = 2.85$, $p = .58$, $I^2 = 0$.

Clinically significant change. Although authors adopted idiosyncratic criteria for clinically significant change across studies, a single index was computed for each study. Generally, the index was identified by the authors of the studies as a measure of high end-state functioning or status as treatment responder, and when a study reported both, the arithmetic mean was used to compute a single measure of clinically significant change.

The weighted average percentage meeting criteria for clinically significant change at posttreatment was 44% for CT and 45% for RT. See Table 2 for data from individual studies. The weighted odds ratio comparing cognitive and relaxation therapies was 1.02, $p = .95$, CI = 0.55, 1.91, indicating that the groups did not differ in terms of relative odds of achieving clinically significant change at posttreatment. There was no evidence of heterogeneity, $Q(4) = 2.36$, $p = .67$, $I^2 = 0$.

Drop-out rates. The weighted average percentage of drop-outs was 15% for CT and 19% for RT (see Table 2). The weighted odds ratio comparing the two therapies was 0.60, $p = .23$, CI = 0.26, 1.39. There was no evidence of heterogeneity, $Q(4) = 2.24$, $p = .69$, $I^2 = 0$.

Summary

The present analyses suggest that overall both CT and RT effected similar change in the areas of generalized anxiety, anxiety-related cognitions, and depression for people with GAD. Between-groups mean difference effect sizes yielded no appreciable differences between CT and RT in any domain, nor did the groups differ in terms of clinically significant change or drop-out rates. That the groups did not differ in terms of drop-out rates mitigates the potential bias introduced by reliance on completer, rather than intention-to-treat, samples.

PD

The outcome literatures evaluating CBT for PD and PDA are rather distinct. PDA is characterized by phobic avoidance, and treatment necessarily includes in vivo exposure to feared situations, no different from other phobias. In contrast, there are not necessarily any particular external phobic stimuli in PD. Rather, prominent in the cognitive model of PD are catastrophic misinterpretations of benign bodily sensations; these are amenable to modification via CT and RT, without in vivo exposure to feared external situations.

A key component in treating panic is interoceptive exposure to misinterpreted bodily sensations such as heart rate, breathing, sweating, and dizziness. There is considerable and strong evidence that treatment packages that include interoceptive exposure are efficacious for panic (e.g., Chambless & Peterman, 2004), and there is some reason to believe that the exposure itself provides important, incremental effects for panic outcomes. Craske, Rowe, Lewin, and Noriega-Dimitri (1997) found that individuals with PDA who were treated with cognitive restructuring, in vivo exposure, and interoceptive exposure were more improved in terms of panic frequency than were others who received the same package but with breathing retraining in lieu of interoceptive exposure. The implications of these results are not entirely consistent with those from Arntz (2002), who compared CT and interoceptive exposure directly and found no group differences on panic symptoms. It is noteworthy that CT constituted a full treatment condition in Arntz (2002) and also included behavioral experiments composed of exposures to bodily sensations, albeit not necessarily long enough for habituation to occur.

In light of the aforementioned distinction between treatment for PD and PDA, only studies that compared CT and RT for panic disorder without severe agoraphobia were included in these analyses. There were no cases of studies for which categorization was unclear; this and other decisions (e.g., classification of variables) were made on the basis of complete agreement between the authors. CT as conducted in these studies differs from that used in studies of GAD. Whereas CT for GAD can be relatively free of behavioral techniques such as exposure or relaxation, all comparisons between CT and RT for PD but one (J. G. Beck, Stanley, Baldwin, Deagle, & Averill, 1994) incorporated interoceptive behavioral experiments or exposure in the cognitive treatment package. CT for PD is more accurately labeled CBT.

Table 2
Generalized Anxiety Disorder Between-Groups Standardized Mean Difference Effect Sizes at Posttreatment and Rates of Clinically Significant Change (CSC) and Drop-Out

				% CSC		% drop-out	
Study	Anxiety	Cognitive	Depression	CT	RT	CT	RT
Arntz (2003)	−0.10		−0.22	38	40	20	15
Barlow et al. (1992)	−0.15	−0.20	0.22	46	60	24	38
Borkovec et al. (2002)	0.15	−0.21	0.08	43	57	8	15
Butler et al. (1991)	0.64	0.72	0.44	32	16	0	14
Öst & Breitholtz (2000)	−0.12	0.17	−0.22	62	53	5	12
Weighted	0.11	0.13	0.05	44	45	15	19

PD studies were located by means similar to those for GAD studies: a PsycINFO search and manual scanning of the same journals and other relevant published literature. Search terms were keywords *cognitive therapy* or *cognitive–behavior therapy* and *panic disorder*. Inclusionary criteria were also similar except that one study (J. G. Beck et al., 1994) was included in which therapy was conducted in small groups of 4 to 6 individuals. This decision was made on the basis of a study by Telch et al. (1993), who found that CBT was as effective when delivered in groups of that size as when delivered individually in other research trials. Five studies were identified that compared CT and RT for PD without agoraphobia (Arntz & van den Hout, 1996; Barlow, Craske, Cerny, & Klosko, 1989; J. G. Beck et al., 1994; Clark et al., 1994; Öst & Westling, 1995). It should be noted that new versions of CBT for PD do not typically incorporate PR or AR (e.g., Craske & Barlow, 2006); hence, CBT as administered in these studies closely reflects the treatment package as recommended by experts in the treatment of PD.

Again, the five studies were methodologically sound. In all studies, participants were screened using validated structured interviews, and assessor ratings were given by independent and blind raters. Therapists in all studies received regular supervision to maintain treatment quality and adherence to protocol, and integrity was formally assessed in three studies and found to be high (Barlow et al., 1989; J. G. Beck et al., 1994; Clark et al., 1994). Perceived treatment credibility or client expectations were assessed in all studies except Arntz and van den Hout (1996) and never differed by treatment group. Study participants were all between 33 and 38 years old, on average, and in all studies therapists were crossed with treatment condition.

It appears from the extant literature that CBT is more efficacious than RT in treating PD, at least for some outcome measures. Higher rates of participants receiving CBT were panic free at posttreatment in all studies, and the same was true of measures of clinically significant change except in Barlow et al. (1989). In a manner similar to that used for GAD studies, data were meta-analyzed using completer samples. Measures were this time categorized in five domains: primary measures of panic, fear of anxiety, and panic-related cognitions, and secondary measures of anxiety and depression. Effect sizes were calculated for posttreatment between-groups differences. Particular measures that composed each domain for each study are given in Table 3. Effect sizes were also calculated to assess differential rates of panic-free status, as well as clinically significant change and drop-outs. Again, clinically significant change includes measures of change and high end-state functioning.

Results

Between-groups comparisons. Between-groups effect sizes were computed from uncontrolled posttreatment means and standard deviations, and are presented in Table 4. Supplementary data to those published in Arntz and van den Hout (1996) and Öst and Westling (1995) were provided by A. Arntz (personal communication, January 27, 2006) and L.-G. Öst (personal communication, April 24, 2006), respectively, to facilitate the analyses.

Participants who received CBT were significantly better off at posttreatment than were those who received RT in all primary domains: They experienced fewer panic symptoms, had less fear of anxiety, and were more improved in terms of panic-related cognitions. The weighted mean differences were as follows: for panic, $g = 0.36$, $p = .02$, CI = 0.05, 0.66; for fear of anxiety, $g = 0.64$, $p < .001$, CI = 0.31, 0.97; for panic-related cognitions, $g = 0.48$, $p = .006$, CI = 0.14, 0.83. Homogeneity tests were nonsignificant and indicated only mild to moderate heterogeneity: for panic, $Q(4) = 6.35$, $p = .17$, $I^2 = 37$; for fear of anxiety, $Q(3) = 3.28$, $p = .35$, $I^2 = 9$; and for panic-related cognitions, $Q(3) = 4.53$, $p = .21$, $I^2 = 34$.

The CBT and RT groups, however, did not differ in terms of secondary measures of generalized anxiety or depression, for

Table 3
Measures Included in Panic Disorder Without Agoraphobia Meta-Analysis

Study	Panic	Fear of Anxiety	Cognitive	Anxiety	Depression
Arntz & van den Hout (1996)	Panic frequency	Fear of Fear (van den Hout et al., 1987), FQ-Agoraphobia Fear		SCL-90 Anxiety, STAI-T	Bouman, SCL-90 Depression
Barlow et al. (1989)	ADIS-R Severity, number of panic attacks, panic intensity		CSAQ-Cognitive	CSAQ-Somatic, HARS, STAI-T, average daily anxiety	BDI, HDRS, daily depression
J. G. Beck et al. (1994)	ADIS-R Severity, panic frequency	ASI, FQ-Agoraphobia, worry about panic	ACQ	HARS, STAI-T	HDRS
Clark et al. (1994)	Panic frequency (assessor and self), panic distress (assessor and self)	BSQ, FQ-Agoraphobia Fear	ACQ, BSIQ	BAI, HARS (without panic), general tension and anxiety (assessor and self)	BDI
Öst & Westling (1995)		BSQ		BAI, HARS, STAI-T, SAS	BDI, HDRS

Note. FQ = Fear Questionnaire (Marks & Mathews, 1979); SCL-90 Anxiety = Symptom Checklist-90 Anxiety—Dutch version (Arrindell & Ettema, 1981); STAI-T = State-Trait Anxiety Inventory—Trait scale (Spielberger et al., 1970); Bouman = Bouman Depression Inventory (Bouman, 1987); ADIS-R = Anxiety Disorders Interview Schedule—Revised (DiNardo & Barlow, 1988); CSAQ = Cognitive Somatic Anxiety Questionnaire (Schwartz et al., 1978); HARS = Hamilton Anxiety Scale (Hamilton, 1959); BDI = Beck Depression Inventory (A. T. Beck et al., 1961); HDRS = Hamilton Depression Rating Scale (Hamilton, 1960); ASI = Anxiety Sensitivity Inventory (Reiss et al., 1986); ACQ = Agoraphobic Cognitions Questionnaire (Chambless et al., 1984); BSQ = Bodily Sensations Questionnaire (Chambless et al., 1984); BSIQ = Body Sensations Interpretation Questionnaire—panic scale (Clark et al., 1997); BAI = Beck Anxiety Inventory (A. T. Beck et al., 1988); SAS = Self-Rating of Anxiety Scale (Zung, 1971).

Table 4
Panic Disorder Without Agoraphobia Between-Groups Standardized Mean Difference Effect Sizes at Posttreatment and Rates of Panic Free Status, Clinically Significant Change (CSC), and Drop-Out

Study	Panic	Fear of anxiety	Cognitive	Anxiety	Depression	% panic free		% CSC		% drop-out	
						CBT	RT	CBT	RT	CBT	RT
Arntz & van den Hout (1996)	0.38	0.52		0.21	0.15	78	50			0	5
Barlow et al. (1989)	−0.20		0.10	−0.48	−0.57	85	60	52	67	6	33
J. G. Beck et al. (1994)	0.09	0.20	0.30	−0.62	−0.23	65	47	82	68	23	5
Clark et al. (1994)	1.02	0.92	1.08	0.52	0.57	90	50	80	25	5	5
Öst & Westling (1995)	0.32	0.95	0.34	0.26	−0.03	74	65	74	47	0	11
Weighted	0.36	0.64	0.48	0.03	0.03	77	53	72	50	12	14

which the small and nonsignificant weighted mean effect sizes were as follows: for generalized anxiety, $g = 0.03$, $p = .86$, CI = −0.28, 0.33; for depression, $g = 0.03$, $p = .85$, CI = −0.27, 0.33. Tests of homogeneity were as follows: for generalized anxiety, $Q(4) = 8.16$, $p = .09$, $I^2 = 51$, and for depression, $Q(4) = 5.58$, $p = .23$, $I^2 = 28$. Hence, Q was significant for generalized anxiety at $p < .10$, indicating a random-effects model. This was the only homogeneity test, however, that reached statistical significance, and the between-groups effect size was approximately 0. Thus, the results of any random-effects analysis would not change the interpretation of the data, and in the interest of consistency, the results of fixed-effects analyses are reported.

Panic-free status. Across studies, the weighted proportions of participants who were panic free posttreatment were 77% in the CBT conditions and 53% in the RT conditions (see Table 4). A significantly higher proportion of individuals who received CBT were panic free posttreatment compared with those receiving RT. The odds ratio effect size comparing the groups was 2.94, $p = .002$, CI = 1.48, 5.83. There was no heterogeneity, $Q(4) = 2.86$, $p = .58$, $I^2 = 0$.

Clinically significant change. Clinically significant change for each study was computed in a manner similar to that used in the GAD analysis. This measure did not include number of participants panic free.

At posttreatment, the weighted proportions of participants achieving clinically significant change were 72% for CBT and 50% for RT (see Table 4). A greater percentage of participants who received CBT were categorized as having achieved clinically significant change at posttest compared with those who were given RT. The odds ratio effect size for the group difference between CBT and relaxation was 3.22, $p = .004$, CI = 1.46, 7.09, and there was moderate but nonsignificant heterogeneity, $Q(3) = 6.11$, $p = .11$, $I^2 = 51$.

Drop-out rates. The weighted proportions of drop-outs were 12% and 14% from the CBT and RT conditions, respectively (see Table 4). The odds ratio effect size was 0.63, $p = .45$, CI = 0.19, 2.08. An analysis of homogeneity yielded $Q(4) = 6.27$, $p = .18$, $I^2 = 36$.

Summary

The results of the present analyses demonstrate that for PD, CBT was superior to RT in all panic-related domains, including measures of panic symptoms, fear of anxiety, and panic-related cognitions. CBT groups further achieved higher rates of status as panic free and as having experienced clinically significant change than did RT groups, and these rates were high (77% panic free; 72% clinically significant change). The groups did not differ, however, in terms of generalized anxiety or depression. They also did not differ in terms of drop-out rates, suggesting that reliance on completer samples did not affect the results notably.

Discussion

This study had two primary objectives: (a) to address claims that among bona fide treatments, no one is more efficacious than another, and (b) to compare the relative efficacy of CT and RT in the treatment of GAD and PD.

CT and RT appeared equally efficacious in treating GAD, across all domains. The treatments did not differ in terms of anxiety, anxiety-related cognitions, depression, clinically significant change, or drop-out rates. Nevertheless, rates of clinically significant change were modest in both conditions relative to those caused by cognitive–behavioral treatments for other anxiety disorders (e.g., Roemer & Orsillo, 2002), and further research is necessary to improve our understanding of GAD and ability to treat it. Although it is possible that individuals with GAD respond equally well to multiple treatments, the disappointing rates of clinical change are consistent with the possibility that researchers have yet to identify a core feature of the psychopathology and how to treat it. A number of promising avenues are currently being investigated, including the roles of interpersonal problems (e.g., Borkovec et al., 2002), emotional dysregulation (e.g., Mennin, Heimberg, Turk, & Fresco, 2005), and metacognitions (e.g., Ladouceur et al., 2000; Wells, 1995), as well as the possibility that schema-based therapy is appropriate given the traitlike features of GAD (Arntz, 2003; Arntz & Weertman, 1999). At this point, with no clear reason to recommend CT over RT or vice versa, therapists should allow patient preference to guide treatment recommendation.

In contrast, CBT was superior to RT in treating PD for all domains directly relevant to panic: CBT outperformed RT in the domains of panic, fear of anxiety, panic-related cognitions, percent panic free, and clinically significant change. Cast as a binomial effect size display (Rosenthal & Rubin, 1982), the effect size for panic is equivalent to an increase in treatment success rate from

41% to 59% for CT over RT. For fear of anxiety, the increase is from 35% to 65%, and for panic-related cognitions, from 38% to 62%. Moreover, rates of clinically significant change and participants who achieved panic-free status were objectively high for those who completed CBT. In light of these findings as well as previous meta-analyses suggesting that CBT is superior to pharmacotherapy in the treatment of PD (e.g., Gould et al., 1995; Mitte, 2005b), CBT remains the first line of treatment for PD.

One limitation with respect to the generalizability of these findings is the use of a fixed-effects model. There was little evidence of heterogeneity: I^2 values indicated degrees of heterogeneity ranging from absent to medium, and only one analysis of homogeneity (for a domain in which the treatments did not differ) was significant at $\alpha = .10$. Hence, fixed-effects analyses were justified and appropriate to maximize power. Nevertheless, the implications are limited to the pool of studies included in the analyses. Another potential limitation is the possibility of publication bias; however, it is unlikely that the results of large-scale randomized clinical trials would remain unpublished, especially considering that for studies comparing CT and RT, a lack of significant group differences is meaningful. Finally, the results of this study do not address the mechanisms of action of these treatments for these disorders. Although CT more directly targets cognitive change, there is considerable evidence that numerous treatment techniques, when successful in treating anxiety, induce cognitive shifts, including psychotropic medication and exposure without CT (e.g., Chambless & Peterman, 2004; Feske & Chambless, 1995).

These studies illustrate a number of issues that complicate interpretation. First, the studies used inconsistent and idiosyncratic measures of clinically significant change, both in terms of magnitude and type of marker of improvement. To illustrate, in the GAD studies, criteria varied from absolute cutoff scores on particular measures (e.g., Butler et al., 1991) to percentage improvement (e.g., Barlow et al., 1992), to criteria based on the standardized distance from the mean of the patient group (e.g., Öst & Breitholtz, 2002). Second, follow-up data are often difficult to synthesize for at least two reasons: There is no consistent time period at which researchers conduct follow-up assessments, and participants are able to pursue further treatment after the posttreatment assessment. In terms of the former, in neither the GAD nor PD sets of studies did more than three studies report follow-up data for any given time point. This is no doubt complicated by the fact that there is no obvious point at which it is theoretically interesting to assess outcome. In addition, pragmatic complications such as cost and attrition can be insurmountable for researchers. Researchers should routinely assess type and amount of further treatment received, however, and include the data in follow-up analyses by comparing groups or controlling for subsequent treatment. These studies did not do so consistently. For all these reasons, effect sizes for follow-up data were not calculated. Third, only Borkovec et al. (2002) and Butler et al. (1991) assessed an important common factor, the quality of the therapeutic relationship; neither found group differences. In the future it is important that researchers include measures that can help evaluate directly whether apparent differences caused by active ingredients are attributable to common factors such as the therapeutic relationship.

A number of methodological strengths are evident from these studies as well, some of which are relevant to the potential role of common factors. Therapists in all studies were crossed with treatment condition, thereby militating against the likelihood that therapist effects influenced treatment outcome. Furthermore, all authors except Arntz and colleagues (Arntz, 2003; Arntz & van den Hout, 1996) assessed the common factors of perceived treatment credibility and client expectations, which were high and never differed by treatment group. Also worthy of note, observational data were all collected from raters blind to treatment condition.

This meta-analysis supports the utility of research efforts that aim to improve psychotherapy for specific disorders by identifying and improving active ingredients relevant to the study population. Some researchers have claimed that treatment effects are due to common factors rather than active ingredients, and that no two bona fide treatments are differentially efficacious for treating a disorder (e.g., Luborsky et al., 2002; Wampold et al., 1997). Critics have argued, however, that such results are based on overgeneralizations from the finding that all treatments for all disorders when combined do not differ from one another, to the conclusion that specific treatments for specific disorders do not differ (e.g., Chambless, 2002). This article provides a demonstration of that rebuttal. There is ample evidence that both CT and RT qualify as bona fide treatments for both GAD and PD. Moreover, GAD and PD are phenomenologically and epidemiologically similar and highly comorbid (e.g., Brown & Barlow, 1992). That CT and RT do not differ in the treatment of GAD, but do for PD, is evidence for the specificity of treatment to disorder, even for two treatments within a CBT class, and two disorders within an anxiety class. This is even more dramatically evident considering that three research groups were responsible for studies of both GAD and PD. This is not to say that specifically efficacious treatments will be identified for all possible presenting problems. Additional research synthesis of the type conducted here would help to answer this question.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48,* 251–257.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Arntz, A. (2002). Cognitive therapy versus interoceptive exposure as treatment of panic disorder without agoraphobia. *Behaviour Research and Therapy, 40,* 325–341.

*Arntz, A. (2003). Cognitive therapy versus applied relaxation as treatment of generalized anxiety disorder. *Behaviour Research and Therapy, 41,* 633–646.

*Arntz, A., & van den Hout, M. (1996). Psychological treatments of panic disorder without agoraphobia: Cognitive therapy versus applied relaxation. *Behaviour Research and Therapy, 34,* 113–121.

Arntz, A., & Weertman, A. (1999). Treatment of childhood memories: Theory and practice. *Behaviour Research and Therapy, 37,* 715–740.

Arrindell, W. A., & Ettema, H. (1981). Dimensional structure, reliability and validity of the Dutch version of the Symptom Checklist (SCL–90). *Nederlands Tijdschrift voor de Psychologie, 36,* 77–108.

*Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy, 20,* 261–282.

*Barlow, D. H., Rapee, R. M., & Brown, T. A. (1992). Behavioral

treatment of generalized anxiety disorder. *Behavior Therapy, 23,* 551–570.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). *Anxiety disorders and phobias: A cognitive perspective.* New York: Basic Books.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety. *Journal of Consulting and Clinical Psychology, 56,* 893–897.

Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 41,* 561–571.

*Beck, J. G., Stanley, M. A., Baldwin, L. E., Deagle, E. A., III, & Averill, P. M. (1994). Comparison of cognitive therapy and relaxation training for panic disorder. *Journal of Consulting and Clinical Psychology, 62,* 818–826.

Bernstein, D. A., & Borkovec, T. D. (1973). *Progressive relaxation training: A manual for the helping professions.* Champaign, IL: Research Press.

Borkovec, T. D., & Mathews, A. M. (1988). Treatment of nonphobic anxiety disorders: A comparison of nondirective, cognitive, and coping desensitization therapy. *Journal of Consulting and Clinical Psychology, 56,* 877–884.

Borkovec, T. D., & Newman, M. G. (1998). Worry and generalized anxiety disorder. In P. Salkovskis (Ed.), *Adults: Clinical formulation and treatment* (pp. 439–459). Oxford, England: Pergamon.

*Borkovec, T. D., Newman, M. G., Pincus, A. L., & Lytle, R. (2002). A component analysis of cognitive–behavioral therapy for generalized anxiety disorder and the role of interpersonal problems. *Journal of Consulting and Clinical Psychology, 70,* 288–298.

Borkovec, T. D., & Ruscio, A. M. (2001). Psychotherapy for generalized anxiety disorder. *Journal of Clinical Psychiatry, 62,* 37–42.

Borkovec, T. D., & Whisman, M. A. (1996). Psychosocial treatment for generalized anxiety disorder. In M. R. Mavissakalian & P. F. Prien (Eds.), *Long-term treatments of anxiety disorders* (pp. 171–199). Washington, DC: American Psychiatric Association.

Bouman, T. K. (1987). *The measurement of depression with questionnaires.* Unpublished doctoral dissertation, University of Groningen, The Netherlands.

Brown, T. A., & Barlow, D. H. (1992). Comorbidity among anxiety disorders: Implications for treatment and *DSM–IV*. *Journal of Consulting and Clinical Psychology, 60,* 835–844.

Brown, T. A., Barlow, D. H., & Liebowitz, M. R. (1994). The empirical basis of generalized anxiety disorder. *American Journal of Psychiatry, 151,* 1272–1280.

*Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). Comparison of behavior therapy and cognitive behavior therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 59,* 167–175.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 1,* 51–62.

Chambless, D. L. (2002). Beware the dodo bird: The dangers of overgeneralization. *Clinical Psychology: Science and Practice, 9,* 13–16.

Chambless, D. L., Caputo, G. C., Bright, P., & Gallagher, R. (1984). Assessment of fear of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *Journal of Consulting and Clinical Psychology, 52,* 1090–1097.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52,* 685–716.

Chambless, D. L., & Peterman, M. (2004). Evidence on cognitive–behavioral therapy for generalized anxiety disorder and panic disorder: The second decade. In R. L. Leahy (Ed.), *Contemporary cognitive therapy: Theory, research, and practice* (pp. 86–115). New York: Guilford Press.

Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy, 24,* 461–470.

Clark, D. M. (1996). Panic disorder: From theory to therapy. In P. M. Salkovskis (Ed.), *Frontiers of cognitive therapy* (pp. 318–344). New York: Guilford Press.

*Clark, D. M., Salkovskis, P. M., Hackmann, A., Middleton, H., Anastasiades, P., & Gelder, M. (1994). A comparison of cognitive therapy, applied relaxation, and imipramine in the treatment of panic disorder. *British Journal of Psychiatry, 164,* 759–769.

Clark, D. M., Salkovskis, P. M., Öst, L.-G., Breitholtz, E., Koehler, K. A., Westling, et al. (1997). Misinterpretation of body sensations in panic disorder. *Journal of Consulting and Clinical Psychology, 65,* 203–213.

Craske, M. G., & Barlow, D. H. (2006). *Mastery of your anxiety and panic: Therapist guide* (4th ed.). New York: Oxford University Press.

Craske, M. G., Rowe, M., Lewin, M., & Noriega-Dimitri, R. (1997). Interoceptive exposure versus breathing retraining within cognitive–behavioural therapy for panic disorder with agoraphobia. *British Journal of Clinical Psychology, 36,* 85–99.

DiNardo, P., & Barlow, D. H. (1988). *Anxiety Disorders Interview Schedule—Revised (ADIS–R*). Albany, NY: Center for Stress and Anxiety Disorders.

Dugas, M. J., & Koerner, N. (2005). Cognitive–behavioral treatment for generalized anxiety disorder: Current status and future directions. *Journal of Cognitive Psychotherapy, 19,* 61–81.

Feske, U., & Chambless, D. L. (1995). Cognitive behavioral versus exposure only treatment for social phobia: A meta-analysis. *Behavior Therapy, 26,* 695–720.

Gould, R. A., Otto, M. W., & Pollack, M. H. (1995). A meta-analysis of treatment outcome for panic disorder. *Clinical Psychology Review, 15,* 819–844.

Gould, R. A., Otto, M. W., Pollack, M. P., & Yap, L. (1997). Cognitive–behavioral and pharmacological treatment of generalized anxiety disorder: A preliminary meta analysis. *Behavior Therapy, 28,* 285–305.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32,* 50–55.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry, 23,* 56–61.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107–128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558.

Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods, 11,* 193–206.

Katon, W. (1996). Panic disorder: Relationship to high medical utilization, unexplained physical symptoms, and medical costs. *Journal of Clinical Psychiatry, 57,* 11–18.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., et al. (1994). Lifetime and 12-month prevalence of *DSM–III–R* psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19.

Ladouceur, R., Dugas, M. J., Freeston, M. H., Leger, E., Gagnon, F., & Thibodeau, N. (2000). Efficacy of a new cognitive–behavioral treatment for generalized anxiety disorder: Evaluation in a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 68,* 957–964.

Luborsky, L., Rosenthal, R., Diguer, 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 and Practice, 9,* 2–12.

Marin-Martinez, F., & Sanchez-Meca, J. (1999). Averaging dependent effect sizes in meta-analysis: A cautionary note about procedures. *Spanish Journal of Psychology, 2,* 32–38.

Markowitz, J. S., Weissman, M. M., Ouellette, R., Lish, J. D., & Klerman, G. L. (1989). Quality of life in panic disorder. *Archives of General Psychiatry, 46,* 984–992.

Marks, I. M. (1987). *Fears, phobias, and rituals: Panic, anxiety, and their disorders.* New York: Oxford University Press.

Marks, I. M., & Mathews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy, 17,* 263–267.

Mennin, D. S., Heimberg, R. G., Turk, C. L., & Fresco, D. M. (2005). Preliminary evidence for an emotion dysregulation model of generalized anxiety disorder. *Behaviour Research and Therapy, 43,* 1281–1310.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28,* 487–495.

Mitte, K. (2005a). Meta-analysis of cognitive–behavioral treatments for generalized anxiety disorder: A comparison with pharmacotherapy. *Psychological Bulletin, 131,* 785–795.

Mitte, K. (2005b). A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *Journal of Affective Disorders, 88,* 27–45.

Ohman, A., & Soares, J. J. F. (1994). "Unconscious anxiety": Phobic responses to masked stimuli. *Journal of Abnormal Psychology, 103,* 231–240.

Öst, L.-G. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research and Therapy, 25,* 397–409.

Öst, L.-G. (1988). Applied relaxation vs progressive relaxation in the treatment of panic disorder. *Behaviour Research and Therapy, 26,* 13–22.

Öst, L.-G. (1992). *Applied relaxation: Manual for a behavioral coping technique.* Unpublished manuscript, Department of Psychology, Stockholm University.

*Öst, L.-G., & Breitholtz, E. (2000). Applied relaxation vs. cognitive therapy in the treatment of generalized anxiety disorder. *Behaviour Research and Therapy, 38,* 777–790.

Öst, L.-G., Thulin, U., & Ramnerö, J. (2004). Cognitive behavior therapy vs. exposure in vivo in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy, 42,* 1105–1127.

*Öst, L.-G., & Westling, B. E. (1995). Applied relaxation vs cognitive behavior therapy in the treatment of panic disorder. *Behaviour Research and Therapy, 33,* 145–158.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency, and individual differences in fearfulness. *Behaviour Research and Therapy, 24,* 1–8.

Roemer, L., & Orsillo, S. M. (2002). Expanding our conceptualization of and treatment for generalized anxiety disorder: Integrating mindfulness/acceptance-based approaches with existing cognitive–behavioral models. *Clinical Psychology: Science and Practice, 9,* 54–68.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74,* 166–169.

Sanderson, W. C., & Wetzler, S. (1991). Chronic anxiety and generalized anxiety disorder: Issues in comorbidity. In R. M. Rapee & D. H. Barlow (Eds.), *Chronic anxiety: Generalized anxiety disorder and mixed anxiety-depression* (pp. 119–135). New York: Guilford Press.

Schwartz, G. E., Davidson, R. J., & Goleman, D. J. (1978). Patterning of cognitive and somatic processes in the self-regulation of anxiety: Effects of meditation versus exercise. *Psychosomatic Medicine, 40,* 321–328.

Snaith, R. P., Bridge, C. W. K., & Hamilton, M. (1976). The Leeds scales for the self-assessment of anxiety and depression. *British Journal of Psychiatry, 128,* 156–165.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual of the State–Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Telch, M. J., Lucas, J. A., Schmidt, N. B., Hanna, H. H., Jaimez, T. S., & Lucas, R. A. (1993). Group cognitive–behavioral treatment of panic disorder. *Behaviour Research and Therapy, 31,* 279–287.

van den Hout, M. A., van der Molen, G. M., Griez, E., & Lousberg, E. (1987). Specificity of interoceptive fear to panic disorder. *Journal of Psychopathology and Behavioral Assessment, 9,* 99–109.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes." *Psychological Bulletin, 122,* 203–215.

Watson, J. P., & Marks, I. M. (1971). Relevant and irrelevant fear in flooding: A crossover study of phobic patients. *Behavior Therapy, 2,* 275–293.

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist, 61,* 671–689.

Wells, A. (1995). Meta-cognition and worry: A cognitive model of generalized anxiety disorder. *Behavioural and Cognitive Psychotherapy, 23,* 301–320.

Wittchen, H.-U., Zhao, S., Kessler, R. C., & Eaton, W. W. (1994). *DSM–III–R* generalized anxiety disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 355–364.

Zung, W. W. K. (1971). A rating instrument for anxiety disorders. *Psychosomatics, 12,* 371–379.

Received August 23, 2006
Revision received January 3, 2007
Accepted March 18, 2007 ■

Correction to Siev and Chambless (2007)

In the article, "Specificity of Treatment Effects: Cognitive Therapy and Relaxation for Generalized Anxiety and Panic Disorders," by Jedidiah Siev and Dianne L. Chambless (*Journal of Consulting and Clinical Psychology*, 2007, Vol. 75, No. 4, pp. 513–522), the individual measures were not listed in the domains labeled "Panic" and "Cognitive" for the Öst and Westling (1995) citation in Table 3. The corrected table appears below, with the added text appearing in bold font.

Table 3
Measures Included in Panic Disorder Without Agoraphobia Meta-Analysis

Study	Panic	Fear of Anxiety	Cognitive	Anxiety	Depression
Arntz & van den Hout (1996)	Panic frequency	Fear of Fear (van den Hout et al., 1987), FQ-Agoraphobia Fear		SCL-90 Anxiety, STAI-T	Bouman, SCL-90 Depression
Barlow et al. (1989)	ADIS-R Severity, number of panic attacks, panic intensity		CSAQ-Cognitive	CSAQ-Somatic, HARS, STAI-T, average daily anxiety	BDI, HDRS, daily depression
J. G. Beck et al. (1994)	ADIS-R Severity, panic frequency	ASI, FQ-Agoraphobia, worry about panic	ACQ	HARS, STAI-T	HDRS
Clark et al. (1994)	Panic frequency (assessor and self), panic distress (assessor and self)	BSQ, FQ-Agoraphobia Fear	ACQ, BSIQ	BAI, HARS (without panic), general tension and anxiety (assessor and self)	BDI
Öst & Westling (1995)	**Diary, panic frequency, panic distress**	BSQ	**ACQ**	BAI, HARS, STAI-T, SAS	BDI, HDRS

Note. FQ = Fear Questionnaire (Marks & Mathews, 1979); SCL-90 = Symptom Checklist-90—Dutch version (Arrindell & Ettema, 1981); STAI-T = State-Trait Anxiety Inventory—Trait scale (Spielberger et al., 1970); Bouman = Bouman Depression Inventory (Bouman, 1987); ADIS-R = Anxiety Disorders Interview Schedule—Revised (DiNardo & Barlow, 1988); CSAQ = Cognitive Somatic Anxiety Questionnaire (Schwartz et al., 1978); HARS = Hamilton Anxiety Scale (Hamilton, 1959); BDI = Beck Depression Inventory (A. T. Beck et al., 1961); HDRS = Hamilton Depression Rating Scale (Hamilton, 1960); ASI = Anxiety Sensitivity Inventory (Reiss et al., 1986); ACQ = Agoraphobic Cognitions Questionnaire (Chambless et al., 1984); BSQ = Bodily Sensations Questionnaire (Chambless et al., 1984); BSIQ = Body Sensations Interpretation Questionnaire—panic scale (Clark et al., 1997); BAI = Beck Anxiety Inventory (A. T. Beck et al., 1988); SAS = Self-Rating of Anxiety Scale (Zung, 1971).