A Component Analysis of Cognitive–Behavioral Therapy for Generalized Anxiety Disorder and the Role of Interpersonal Problems

T. D. Borkovec, Michelle G. Newman, and Aaron L. Pincus
The Pennsylvania State University

Clients with generalized anxiety disorder (GAD) received either (a) applied relaxation and self-control desensitization, (b) cognitive therapy, or (c) a combination of these methods. Treatment resulted in significant improvement in anxiety and depression that was maintained for 2 years. The large majority no longer met diagnostic criteria; a minority sought further treatment during follow-up. No differences in outcome were found between conditions; review of the GAD therapy literature suggested that this may have been due to strong effects generated by each component condition. Finally, interpersonal difficulties remaining at posttherapy, measured by the Inventory of Interpersonal Problems Circumplex Scales (L. E. Alden, J. S. Wiggins, & A. L. Pincus, 1990) in a subset of clients, were negatively associated with posttherapy and follow-up improvement, suggesting the possible utility of adding interpersonal treatment to cognitive–behavioral therapy to increase therapeutic effectiveness.

Generalized anxiety disorder (GAD) is a common problem (with prevalence rates of 3.6% to 5.1% over a lifetime and 3.1% over 1 year; Kendler, Neale, Kessler, Heath, & Eaves, 1992; Wittchen, Zhao, Kessler, & Eaton, 1994) and is one of the most common comorbid conditions among other anxiety and mood disorders (Brown & Barlow, 1992). GAD tends to be chronic with an early onset and a resistance to change (Sanderson & Wetzler, 1991; Zuellig & Newman, 1996), leading to the suggestion that GAD may be the basic anxiety disorder from which other disorders often emerge (e.g., Brown, Barlow, & Liebowitz, 1994). Finally, its core symptom of worry has been found to relate to significant health problems, including cardiovascular disease (see Brosschot & Thayer, in press; Newman, 2000). For all of these reasons, development of effective psychosocial treatments for GAD is an important priority.

Prior reviews of GAD outcome research have indicated that cognitive–behavioral therapy (CBT) is indeed efficacious, yielding improvements superior to waiting-list–no-treatment conditions as well as to conditions that control for nonspecific or common factors (Borkovec & Ruscio, 2000; Borkovec & Whisman, 1996; Chambless & Gillis, 1993; Gould, Otto, Pollack, & Yap, 1997). Such findings have resulted in CBT being listed as an empirically supported treatment for GAD (Chambless et al., 1998). Moreover, treatment gains from CBT are routinely maintained or increased at long-term follow-up, the therapy is well received by clients in clinical trials (with an average drop-out rate of only 8.3%), and psychotropic medication usage often declines by follow-up (Borkovec & Ruscio, 2000; Hoehn-Saric, Borkovec, & Nemiah, 1995). Despite this relative success, at least three significant and interrelated issues remain. First, it is unclear what the relative contributions are of the cognitive and behavioral components commonly included in CBT packages. Meta-analysis shows that CBT is associated with the largest within-group and between-group effect sizes relative to either behavior therapy alone or cognitive therapy alone. However, direct statistical contrasts within each relevant study have found significant superiority of CBT in only 2 of 10 posttherapy comparisons and 3 of 7 follow-up comparisons (Borkovec & Ruscio, 2000). The latter findings may be related to an issue of power. Further investigation using sufficient sample size to identify significant causative ingredients among CBT techniques individually administered could potentially facilitate our knowledge about the nature of the disorder and the nature of the mechanisms of CBT-induced change and could thus contribute to further therapeutic developments for incrementing its effectiveness. Second, estimates from the GAD literature indicate that only 50% of clinical trial participants meet high endstate functioning criteria (Borkovec & Whisman, 1996). Thus, it is important to pursue ways of increasing the therapeutic effects of our therapies for this disorder. Third, we know very little about individual differences among GAD clients that are predictive of responsiveness to CBT treatment (see Hoehn-Saric, Borkovec, & Crits-Christoph, in press; Newman, 2000). Identification of such factors could provide guidance for making additions to or modifications in CBT that might increase the therapy’s effectiveness.

In our prior GAD investigation (Borkovec & Costello, 1993), applied relaxation training and a CBT package (applied relaxation, self-control desensitization, and brief cognitive therapy) were both found to be superior to a nondirective (supportive listening) control condition on primary anxiety measures at the posttherapy
assessment. Thus they were both demonstrated to contain active ingredients beyond nonspecific or common factors. Although parametric analyses of the primary outcome measures failed to find condition differences at follow-up, CBT was associated with the greatest maintenance of gains and with a significantly superior degree of clinically significant change at 12-month follow-up relative to the other two conditions. This latter outcome suggested that one or more of the therapeutic ingredients contained within our CBT package (cognitive therapy and/or imagery rehearsal of cognitive and relaxation coping responses in self-control desensitization) beyond applied relaxation training and beyond nonspecific factors is potentially causative of better outcome in the long term.

On the basis of the reviews of the GAD treatment literature and the above study, the original goals of the present investigation were twofold. First, attempts to identify critical ingredients in CBT were continued by comparing the package with two of its components: (a) cognitive therapy (CT) alone and (b) applied relaxation training combined with self-control desensitization (SCD) for rehearsing relaxation and rational self-statement coping responses in imagery. Second, an attempt was made to increment the effectiveness of CBT by increasing therapy time by 50% over our prior trial to allow for more thorough cognitive therapy. Given the extensive involvement of numerous cognitive processes in GAD and in the nature of worry (see Borkovec, Alcaine, & Behar, in press), thorough cognitive therapy might well be expected to increase the therapeutic effects of CBT.

Near the end of the 2nd year of this 5-year study, we began to administer the Inventory of Interpersonal Problems—Circumplex Scales (IIP–C; Alden, Wiggins, & Pincus, 1990; Horowitz, Alden, Wiggins, & Pincus, 2000). Our past clinical experience suggested that a subset of GAD clients existed who were hostile in their relationships with others, who blamed others for their anxiety and worry, and who seemed to respond poorly to our interventions. These clients displayed interpersonal characteristics similar to those found by Horowitz, Rosenberg, and Bartholomew (1993) to be resistant to change in psychodynamic therapy. We chose the IIP–C because it assessed these and other interpersonal difficulties and allowed us to evaluate whether a psychometrically sound dimensional assessment of interpersonal functioning would predict outcome.

Because our prior study indicated that the CBT package may yield the greatest long-term outcome and because GAD and its cardinal characteristic of worry involve empirically demonstrated interactions among somatic, cognitive (thought and imagery), and affective domains of functioning (see Borkovec & Newman, 1999; Borkovec, Alcaine, & Behar, in press), we predicted that the CBT treatment targeting all of these domains would generate superior outcome relative to its two component conditions. We also expected that incrementing the amount of therapy time and providing thorough cognitive therapy should result in increased effectiveness relative to prior investigations.

Method

Participants

Of 459 people who contacted our project, 320 were ruled out by phone screening; diagnostic interviews ruled out an additional 63 for not meeting admission criteria. Of 76 admitted clients, 7 dropped out at early stages of treatment (4 in SCD, 2 in CT, and 1 in CBT; nonsignificant by chi-square analysis). No clients were removed for deteriorating conditions during therapy. The 69 clients who completed treatment averaged 37.14 years of age (SD = 11.71), and duration of the GAD problem averaged 12.81 years (SD = 12.07). Ethnicity was represented by 62 White, 2 African American, 3 Hispanic, and 2 Middle Eastern clients. Forty-five clients were women. Nine clients (3 in each condition) had been referred by mental health practitioners; the remainder had responded to media advertisements. Only 2 clients (1 in SCD, 1 in CBT) were taking psychotropic drugs for anxiety; they agreed to maintain dosage and frequency during therapy with their physician’s approval. All of these characteristics were nearly equally distributed among conditions and were not significantly different.

Procedure

Selection and assessor outcome ratings. Clients were enrolled over a 5-year period. Clinical assessors (advanced clinical graduate students trained in DSM diagnostic interviewing) conducted a 30-min phone interview to determine likely diagnostic suitability. They then administered in person a modified version of the Anxiety Disorders Interview Schedule—III–R (ADIS-R; DiNardo & Barlow, 1988), which included the Hamilton Anxiety Rating Scale (HARS; Hamilton, 1959), the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960), Assessor severity of GAD Anxiety Symptoms rating (on a 0–8-point scale), and additional questions in the GAD section corresponding to two criteria being proposed at the time of study initiation by the DSM GAD subcommittee (i.e., uncontrollable worrying, and three of six somatic symptoms). Because GAD is characterized by the lowest degree of interrater reliability among the anxiety disorders (Barlow & DiNardo, 1991), a second ADIS to reduce the likelihood of false positive cases was given within 2 weeks by the therapist who would see the client in therapy upon acceptance into the trial. Each of the four therapists ruled out one or more clients at this point (for a total of 9). Admission criteria included agreement between these two diagnostic interviewers on a principal diagnosis of Diagnostic and Statistical Manual of Mental Disorders (3rd ed., rev.; DSM–III–R; American Psychiatric Association, 1987) GAD, no diagnosable panic disorder (as recommended by the funding agency’s review committee), Assessor Severity rating of 4 (moderate) or greater, absence of concurrent psychosocial therapy, no history of having received actual CBT methods in prior therapy, no medical contributions to the anxiety, no antidepressant medication, and absence of severe depression, substance abuse, psychosis, and organic brain syndrome. All but two clients (97.1%) met both DSM–III–R and DSM–IV (4th ed.; American Psychiatric Association, 1994) GAD criteria. A briefer version of the ADIS (assessing only those diagnoses identified at pretherapy) and the rating scales were readministered 10–14 days after the last therapy session and at 6- and 12-month follow-up assessments; the complete ADIS and rating scales were given at 24-month follow-up. The same assessor administering the preassessment to a client also administered the postassessment to that client; this was the case at follow-up whenever possible. Assessors were kept unaware of condition by preventing their access to the client during the therapy period and to session tape recordings or any other information pertaining to the client’s condition status.

Self-report outcome measures. Clients completed a daily diary four times a day during the 2 weeks before therapy, during therapy (including the 10–14 days after Session 14 for postassessment), and for 1 week before each follow-up. In the diary, clients rated their average level of anxiety during the preceding time block on a 100-point scale (see Barlow et al., 1984).

At a separate questionnaire session, clients completed the State–Trait Anxiety Inventory—Trait version (STAI–T; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983); the Reactions to Relaxation and Arousal Questionnaire (RRAQ; Heide & Borkovec, 1983), a factor-analytically derived measure of fear of relaxation; the Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990), a measure of worry
that distinguishes GAD from all other anxiety disorders (Brown, Antony, & Barlow, 1992); and the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). These measures were given again at the postassessment and follow-up ADIS sessions. Random assignment to conditions occurred within each wave of three clients; assignment to therapist was random within restraints of availability and caseload.

The IIP–C was added to the assessment battery near the end of the 2nd year of the project. The IIP–C has eight scales (Dominating/Controlling, Vindictive/Self-centered, Cold/Distant, Socially isolated, Nonassertive, Overly accommodating, Self-sacrificing, and Insecure/Needy) that form a circumplex of interpersonal problems around the octants of dominance and nurturance. This measure shows strong convergence between self-, therapist-, and peer-rating profiles, discriminates subgroups of depressed and of socially anxious clients, has strong test–retest reliability (total $r = .98$; average subscale $r = .81$), and good alpha ($\beta = .72–.85$) coefficients (Alden & Phillips, 1990; Bartholomew, 1989; Horowitz, Rosenberg, Baer, Ureno, & Villasenor, 1988; Kachin, Newman, & Pincus, 2001). The IIP–C is sensitive to change in psychotherapy, and client improvement as measured by the IIP–C has agreed with improvement in other client self-report measures, therapists’ judgments, and independent observer ratings (Horowitz et al., 1988). Of the 69 completing clients, 43 provided pretherapy IIP–C data and 38 completed posttherapy IIP–C data.

**During-session process measures.** Clients completed a three-item credibility scale (9-point scales) and a 0%–100% Expectancy of Improvement Scale (Borkovec & Mathews, 1988) at the end of the first therapy session. To ensure that treatments did not differ on therapeutic relationship qualities, two measures were administered. The Relationship Inventory (Barrett-Lennard, 1986) assessed client perception of therapist warmth, empathy, unconditionality, and congruence and was given at the end of Sessions 1, 4, 8, and 12. The Working Alliance Inventory (Horvath & Greenberg, 1989) evaluated perceived agreement on the goals and tasks of therapy and strength of bonds as rated separately by both clients and therapists after Sessions 2, 5, 10, and 14. Client forms were sealed in an envelope on completion; they were told that the therapist would never see the completed forms.

**Therapists.** Three doctoral-level therapists (two female and one male) and one advanced clinical graduate student (female) conducted therapy. One of the therapists participated in all 5 years of the trial and saw 34 of the clients; the two female doctoral-level therapists saw 16 and 13 clients, respectively; the graduate student therapist conducted treatment with 6 clients. The three doctoral-level therapists saw a nearly equal number of clients in each condition, and the student therapist saw 2 clients in each condition. The male therapist had served in a prior GAD study (Borkovec & Costello, 1993). All therapists underwent protocol training before study initiation, including instruction and demonstrations by T. D. Borkovec, viewing of therapy tapes, and role playing, all following detailed protocol manuals. The three therapists who were new to GAD protocol treated a pilot client with combined CBT before seeing their first study client. T. D. Borkovec listened to audiotapes of sessions and provided weekly individual supervision to guarantee protocol adherence and to maximize therapy quality.

**Therapy conditions.** Fourteen weekly sessions were administered, with 1 fading session after postassessment. In all conditions, the first 4 sessions were 2.0 hr in duration; remaining sessions were 1.5 hr.

The first 30 min of each SCD and CT session involved only supportive listening, during which clients were told that therapy would partly involve the exploration of life experiences in a relaxed atmosphere, with the goals of facilitating and deepening knowledge about the self and anxiety (see Borkovec & Costello, 1993). The methodological purpose of this portion in CT and SCD was to hold constant the total amount of treatment time while also holding constant the total amount of time devoted to CT in both the CT and CBT conditions and the total amount of time devoted to SCD techniques in both the SCD and CBT. The supportive listening manual instructed therapists to provide an accepting, nonjudgmental, empathic environment; to direct client attention to primary feelings; and to facilitate allowing and acceptance of affect, using supportive statements, reflective listening, and empathic communications. Direct suggestions, advice, interpretative reflections, or use of any CBT methods were prohibited.

Several elements were common to the three conditions, although their content differed according to condition: presentation of a model of anxiety and rationale for therapy, self-monitoring and early identification of anxiety cues, homework assignments, and review of homework including the results of daily self-monitoring and technique practice and applications. Cognitive therapy was conducted according to Beck and Emery (1985) and began with a presentation of a cognitive model of anxiety and treatment plus training in self-monitoring and early identification of environmental, somatic, affective, imaginal, and thought (especially worry) cues that trigger or contribute to interpretations of threat. Clients were told that therapy would focus on learning new, less anxiety-provoking ways of perceiving one’s self, the world, and the future. Anxiety was described as a habitual spiral process wherein the cognitive perception of threat led to interacting anxious reactions that included thoughts (especially worry), images, somatic reactions, and affect. Thus, therapy would involve learning to detect the incipient spiral and to substitute alternative, more accurate perspectives. Over sessions, standard cognitive therapy procedures were used. The therapist worked with the client on specifying cognitive predictions, interpretations, beliefs, and assumptions underlying threatening perceptions of external events and internal cues. Therapy techniques included logical analysis, examination of evidence and probabilities, labeling of logical errors, decatastrophizing, and generation of alternative thoughts and beliefs. Socratic method was emphasized throughout therapy. Over sessions, discussions increasingly focused on multiple alternative perspectives for any given situation. Clients were encouraged to apply alternative perspectives early on identification of incipient anxiety during daily living. Information from daily self-monitoring and discussion with clients contributed to identifying crucial thoughts and underlying themes and beliefs. Imagery was used to aid in thought identification (along with role-play and recall of past anxious episodes), but rehearsal of coping responses in imagery was not allowed. Homework emphasized frequent applications of alternative perspectives and behavioral tasks to provide opportunities to test new beliefs and predictions.

Clients in SCD were told that therapy would involve learning new coping techniques for reducing anxiety and worry. As in CT, anxiety was described as a habitual spiral process. Therapy would thus involve self-monitoring of internal reactions and their sequential nature; learning to catch the spiral early and to intervene with a variety of relaxation responses to anxious thoughts, feelings, and images to disrupt anxious spirals and to create new coping habits; learning to focus attention on present-moment experience rather than on mentally created past events or future possibilities; and imaginal rehearsal of coping methods to facilitate fear extinction and coping-response-habit acquisition. Relaxation training over the sessions included the full course of progressive relaxation training, cue controlled, and differential relaxation training as described in Bernstein and Borkovec (1973), slowed diaphragmatic breathing, relaxing imagery, and meditational relaxation. Of most importance, all of these techniques were used in applied-relaxation training (Ost, 1987), wherein clients learned to deploy their relaxation responses frequently throughout the day and in response to any incipient anxiety cues. Practicing these applications occurred within each therapy session as well. The importance of formal relaxation practice twice a day to strengthen the relaxation response and frequent application during the day was emphasized. In the development of alternative self-statements for use in imagery rehearsals, the therapist was allowed only to ask the client the following standard question: “Can you think of an equally true or more likely alternative self-statement that is less anxiety arousing?” The therapist was not allowed to use logical analysis, search for evidence, decatastrophizing, or any other formal CT method to identify alternative self-statements and no mention of underlying beliefs was made. Over sessions, discussions increasingly focused on the flexible
choice of relaxation methods depending on the internal and external circumstances and on discovery-oriented experimentation. Anxiety cue hierarchies for use during formal SCD were constructed from pretherapy ADIS information, daily self-monitoring, and in-session discussion with the client. During the desensitization procedure, after the client was deeply relaxed, external and internal anxiety cues (including incipient worry about topics of current concern) were presented until the client signaled the presence of anxious feelings. The client then continued imagining the external situation while imagining that he or she was deploying the coping responses. At the elimination of anxious feelings, he or she imagined continued coping deployments for 20 s and then turned off all imagery and focused only on the relaxed state for 20 s. Scenes were repeated until the client could no longer generate anxiety or was able to eliminate it rapidly (i.e., within 5–7 s). Homework emphasized frequent applications of relaxation, focus on living in the present moment, and behavioral approach tasks to provide opportunities for the deployment of newly learned coping strategies.

CT contained all of the treatment techniques described above for CT and SCD, except that no supportive listening element was included and perspective shifts created during CT portions of the session were used during SCD rehearsals along with relaxation responses.1

Quality checking. Audiotapes from 20% of the sessions (three tapes for each client randomly selected from Sessions 1–5, 6–10, and 11–14) were checked for protocol adherence by trained clinical graduate students who listened to the entire session and marked every therapist utterance against a checklist of not-allowed and allowed interventions representing all techniques from the behavioral, cognitive, and supportive-listening treatment protocols. Among the 15,945 checked statements, only seven minor breaks in protocol occurred. A minor break was defined as one or two statements that were inappropriate for the administered treatment; a major break would have involved lengthier use of inappropriate techniques. Five of these breaks occurred during supportive-listening portions of CT sessions (three provisions of information and two interpretive reflections). A sixth break occurred in SCD when the therapist provided a single statement that reflected an alternative perspective; the seventh break happened in CBT when an interpretive reflection was used.

Integrity checking. Audiotapes from 20% of the sessions (three tapes for each client randomly selected from Sessions 1–5, 6–10, and 11–14) were checked for protocol adherence by trained clinical graduate students who listened to the entire session and marked every therapist utterance against a checklist of not-allowed and allowed interventions representing all techniques from the behavioral, cognitive, and supportive-listening treatment protocols. Among the 15,945 checked statements, only seven minor breaks in protocol occurred. A minor break was defined as one or two statements that were inappropriate for the administered treatment; a major break would have involved lengthier use of inappropriate techniques. Five of these breaks occurred during supportive-listening portions of CT sessions (three provisions of information and two interpretive reflections). A sixth break occurred in SCD when the therapist provided a single statement that reflected an alternative perspective; the seventh break happened in CBT when an interpretive reflection was used.

Quality checking. Jeremy Safran served as quality checker for the CT portions of the trial conditions. For each client in a condition containing CT (CT and CBT), he listened to the entirety of two randomly selected session tapes (one from Sessions 1–7, the other from Sessions 8–14) and rated the sessions for overall quality, general clinical skills, and CT skills on 0–6-point scales, using Young and Beck’s (1980) Cognitive Therapy Scale. Condition × Therapist analysis of variance (ANOVA) found no significant effects. Overall quality (M 4.77) was close to “very good” on the scale.

Results

Preliminary Analyses

Table 1 presents means and standard deviations on the six anxiety and two depression outcome measures for each condition at each assessment (follow-up values are based on clients who completed those assessments; see below for information on missing data). A two-way multivariate analysis of variance (MANOVA) (Condition × Therapist) on all eight pretherapy measures indicated no significant main or interaction effects involving condition or therapist. Although the conditions were thus equivalent at the start of therapy, all postassessment and follow-up scores were analyzed using multivariate analyses of covariance (MANCOVA, with pretherapy values as covariates) to provide base-free measures of change.

A two-way MANOVA (Condition × Therapist) applied to the expectancy scale and average credibility scales obtained at the end of the first session yielded no significant effects, with moderately high expectancy (M 67.54%) and a high degree of credibility (M 7.58 on the 9-point scales) reported by the total group.

Postassessment Improvement

Between-groups change. Two-way MANCOVAs (Condition × Therapist) conducted on the six postassessment anxiety measures and separately on the two depression measures found no significant effects.

Within-group change. Related-measures t tests were calculated within each condition on each of the eight anxiety and depression outcome measures from pretherapy to posttherapy. These tests indicated significant improvement (after Sime’s, 1986, Bonferroni correction) on all measures for each of the three conditions, with ts (all dfs 21) ranging from 3.30 to 14.31 and ps ranging from .003 to .001.

Clinically significant change. An endstate functioning measure, used to operationalize the degree to which a client showed clinically meaningful gain, was calculated by summing the number of the six anxiety outcome measures on which the client either fell within one standard deviation of the mean of nonanxious normative samples (HARS, STAI-T, RRAQ, and PSWQ) or a score that exceeded a face-valid level of meaningful change when norms were not available, a score of 2 (mild) or less on the 9-point Assessor severity and 20 (“slight anxiety”) or less on diary severity. Any given client’s endstate score could thus range from 0 to 6. Table 2 presents the percentage of clients in each condition who had an endstate score of 4, 5, or 6 (high endstate); low-endstate percentage would be the high-endstate percentage minus 100. Chi-square analysis indicated no significant condition effects on these frequencies.

Follow-Up Improvement

At 6-month follow-up, 2 SCD clients and 1 CT client failed to attend the assessment session. At 12-month follow-up, 2 SCD clients and 2 CT clients did not complete the assessments. At 24-month follow-up, 2 clients in each of the three conditions declined to undergo the assessments, and one additional CT client from whom interview measures were obtained did not complete the questionnaire materials.

Between-groups change. Two-way MANCOVAs (Condition × Therapist) were conducted at each follow-up assessment on the six anxiety measures and separately on the two depression measures. No significant effects involving condition or therapist factors emerged.

Within-group change. t tests with Sime’s Bonferroni correction evaluated change within each therapy condition from pretherapy to each follow-up assessment on each anxiety and depression outcome measure. These tests indicated significant improvements on 71 of the 72 comparisons, with ts (with dfs ranging from 18–21) ranging from 2.21 to 12.19 and ps ranging

---

1 The CBT protocol therapy manual is available from T. D. Borkovec on request.
from .04 to .001. The one exception involved nonsignificant improvement for CT on the HRSD at 24-month follow-up. Similar within-condition tests of change from posttherapy to the 24-month follow-up found no significant effects of time on any measure; thus, improvements experienced in each condition at the end of therapy were maintained 2 years later.

Clinically significant change: Endstate functioning. The same definition of endstate functioning was applied to follow-up anxiety outcome measures. Table 2 presents the percentage of clients in each condition meeting criteria for high endstate at each follow-up. Chi-square analyses at each period revealed no significant condition effects. Although the number of clients showing high endstate was somewhat higher at posttherapy and 6-month follow-up than at 12- and 24-month follow-up (52.17%, 57.58%, 40.00%, and 46.03%, respectively), McNemar’s change test (Siegel & Castellan, 1988) indicated that this trend was not significant, \( \chi^2(1, N = 62) = 1.19, p < .20. \)

2 All of the main outcome analyses (MANCOVAs on the anxiety and depression measures and chi-square tests on endstate functioning) were repeated using (a) endpoint scores, wherein the last assessment scores of clients completing treatment was used to replace any missing data at follow-up assessments and (b) intent-to-treat scores, wherein the endpoint scores described above were used, and any clients dropping out of treatment before posttherapy assessment received postassessment and follow-up scores using their pretherapy values. Moreover, principal-components factor analysis with oblique rotation was also applied to the eight covariance-adjusted anxiety and depression outcome scores among completers at each posttherapy and follow-up assessment (each analysis resulting in a single factor with eigenvalue greater than 1.00), and the resulting factor scores at each assessment period were submitted to condition by therapist ANOVAs. With one exception, the same pattern of results described above for completers without missing data emerged from all of these analyses (i.e., there were no condition effects at any assessment moment). The exception involved a Condition \( \times \) Therapist interaction for 6-month endpoint depression scores. The one therapist who saw only two
Clinically Significant Change: Within-Group Effect Sizes at Posttherapy and Follow-Ups

Definitions of endstate functioning have varied across prior GAD therapy research studies, with different measures used and different criteria applied to those measures to determine low- and high-endstate categorizations. Consequently, we also calculated within-group effect sizes (posttherapy or follow-up) mean minus pretherapy mean divided by the pretherapy standard deviation) for each condition on each of the three most commonly used measures in the GAD therapy literature (HARS, Assessor Severity, and STAI–T). We then averaged these three effect sizes to compare the degree of change in our conditions with the changes observed in prior investigations using either CBT or its components and using a standardized change score that was based on identical measures. Table 3 presents these effect sizes for each of our conditions at each assessment period along with the average effect sizes at posttherapy and at follow-up (average follow-up duration was 9 months) for the 13 CBT conditions and the 10 behavior therapy only (relaxation or anxiety management) or CT only component conditions from 11 prior GAD therapy investigations that had used all three of these measures (see Borkovec & Ruscio, 2000, for the review and meta-analysis from which the latter values were obtained). Inspection of these values indicates that (a) our CBT and CT conditions generated somewhat greater posttherapy change but largely equivalent follow-up change, relative to the average outcomes of prior CBT conditions, (b) our SCD condition produced changes largely equivalent to prior CBT conditions at posttherapy and follow-up, and (c) both of our component conditions (CT and SCD) yielded considerably greater change than previously investigated component conditions involving CT alone or behavior therapy alone.

Between-group effect sizes (the difference between posttherapy [or follow-up] means of each pair of the study’s three conditions divided by their pooled posttherapy standard deviation) using the same three common measures were also calculated. Average effect sizes at the four assessment periods were 0.12, 0.21, 0.11, and 0.07. Which condition showed the greatest effect size and which the least varied unsystematically among the three conditions over the different periods.

Diagnostic Status

Table 4 presents the percentage of clients in each condition who met GAD diagnostic criteria at the posttherapy and follow-up assessments; the large majority of clients no longer met GAD criteria at these assessments. Chi-square analyses found no significant differences among conditions at any assessment. Although the percentage of clients diagnosed with GAD increased from around 9% at posttherapy to 17% at 24-month follow-up, this increase was not significant by McNemar’s change test, \( \chi^2(1, N = 62) = 2.27, p < .20 \).

Subsequent Therapy

At each follow-up interview, clients indicated whether they had received further psychosocial or psychotropic intervention outside of the project since the preceding assessment period. Chi-square analyses indicated no significant differences between conditions at any assessment. Overall, the rate of subsequent treatment for the total group was very low at the 6- and 12-month assessments (4.50% and 7.80%, respectively) but increased to 18.96% by 24-months, significantly so as indicated by McNemar’s change test, \( \chi^2(1, N = 62) = 8.00, p < .01 \).

Therapeutic Relationship Measures

Two-way repeated measures ANOVAs (condition by session) on each of the four factors of the Relationship Inventory administered at Sessions 1, 4, 8, and 12 found no condition effects but did indicate significant main effects of session on regard, \( F(3, 64) = 9.37, p < .001 \); empathy, \( F(3, 64) = 13.39, p < .001 \); and unconditionality, \( F(3, 64) = 3.53, p < .05 \). In each case, the quality of these relationship factors increased over sessions. The same type of analysis applied separately to each of the six Working Alliance measures (ratings of agreement on bonds, tasks, and goals) by both clients and therapists obtained after Sessions 2, 5, 10, and 14 found no significant main effects of condition, significant session effects (indicating increasing strength of the Working Alliance over sessions) for all six measures, \( F_s(3, 64) \) ranging from 4.29 to 14.78, \( p_s \) ranging from less than .01 to .001, and one significant Condition \( \times \) Session interaction on therapist ratings of tasks, \( F(6, 128) = 2.46, p < .05 \). The latter interaction effect reflected the fact that therapists perceived their agreement with clients on the tasks of therapy to be initially higher in SCD and

---

3 For example, when a more conservative definition of high-endstate functioning (meeting criteria on 5–6 measures) was applied to the present data, percentages of clients showing high endstate over the four assessments were considerably lower (33.33%, 31.34%, 29.23%, and 22.22%). These percentages were not significantly different over time by chi square analysis, but CBT did produce significantly greater high endstate by this more conservative definition at the 6-month follow-up (\( M = 47.83\% \)) than did CT (\( M = 27.27\% \)) or SCD (\( M = 18.18\% \)). This significant difference emerged irrespective of using therapy completer, endpoint, or intent-to-treat data.
Credibility and Expectancy

First-session credibility and expectancy measures were correlated with the 0–6-point endstate-functioning measures at posttherapy and follow-up assessments. Expectancy assessed at Session 1 did not correlate significantly with endstate at any assessment point. Credibility correlated significantly with posttherapy scores within each condition found significant improvements after Sime’s Bonferroni correction only on the Cold/Distant scale for the CT condition, $r(10) = 4.62, p < .001$, and on the Nonassertive scale for the SCD condition, $r(10) = 4.50, p < .001$.

Pretherapy and posttherapy IIP–C scores were also correlated with endstate measures from the posttherapy and follow-up assessments. Of the 32 correlations among the eight pretherapy IIP–C octants and the four assessment moments, only 6 (1 at posttherapy, 4 at 6-month, and 1 at 12-month follow-up) were significant and negative. Three of 6 pretherapy IIP–C correlations remained significant after Sime’s Bonferroni correction, each involving 6-month-endstate correlations with Dominating/Controlling, $r(41) = -.455, p < .002$, Intrusive/Needy, $r(41) = -.425, p < .004$, and Vindictive/Self-centered, $r(41) = -.360, p < .018$; greater interpersonal problems reported just before therapy were associated with poorer outcome. Of the 12 correlations among the eight posttherapy IIP–C octants and the four assessment moments, 15 (5 at posttherapy, 7 at 6-month, 1 at 12-month, and 2 at 24-month follow-up) were significant ($p < .05$) and negative. After Sime’s Bonferroni correction, each of these correlations remained significant. Five involved postendstate correlations with Vindictive/Self-centered, $r(36) = -.552, p < .001$; Intrusive/Needy, $r(36) = -.522, p < .001$; Dominating/Controlling, $r(35) = -.511, p < .001$; Overly accommodating, $r(36) = -.383, p < .018$; and Nonassertive, $r(36) = -.351, p < .031$. The other 7 involved 6-month-endstate correlations with Intrusive/Needy, $r(35) = -.561, p < .001$; Vindictive/Self-centered, $r(35) = -.554, p < .001$; Dominating/Controlling, $r(35) = -.516, p < .001$; Nonassertive, $r(35) = -.377, p < .021$; Overly accommodating, $r(35) = -.370, p < .024$; Cold/Distant, $r(35) = -.341, p < .039$; and Self-sacrificing, $r(35) = -.361, p < .028$. Although the correlations between outcome and post-IIP might conceivably be due merely to a relationship between pretherapy interpersonal problem severities and the six pretherapy anxiety severities, only 4 of 48 correlations were significant (average $r = .108, p < .50$, indicating little shared variance). This is consistent with the distinction between symptomatic distress and interpersonal distress initially identified by Horowitz et al. (1988) in their construction of the IIP. Similarly and more recently,
Kachin et al. (2001) found that subgroups of DSM-defined social phobics derived through cluster analysis of the IIP–C did not differ on severity of anxiety symptoms assessed by the ADIS–R. Partial correlations removing the contribution of the pretherapy anxiety scores were conducted, however, resulting in the same five significant relationships for postendstate outcome and in five of the same seven significant relationships for 6-month endstate (Cold/Distant and Overly accommodating were no longer significant). Thus, interpersonal difficulties remaining at the end of therapy were more frequently associated with posttherapy and 6-month follow-up outcomes than were pretherapy interpersonal problems and having vindictive, intrusive, and domineering relationships were particularly predictive of poor immediate and intermediate outcome.4

Discussion

The prediction that CBT would be superior to its components was not supported. Furthermore, our efforts to increase the therapeutic effectiveness of CBT were not successful. On the other hand, preliminary correlational evidence from a subset of clients did tentatively suggest the possible role of interpersonal problems in the maintenance of GAD. Below we elaborate on each of these points.

All outcome analyses indicated that both CT and SCD were as effective as CBT that contained all of these techniques. Previous component control and CBT outcome studies have sometimes demonstrated differences in outcome and sometimes not. As mentioned in the introduction, one possible explanation for these mixed results may reside in the power of a study. In investigations finding evidence for differences between CBT and one of its components (Borkovec et al., 1987; Borkovec & Costello, 1993; Butler, Fennell, Robson, & Gelder, 1991; Durham et al., 1994), number of clients per condition ranged from 14 to 20 ($M = 17.25$). In two of the three investigations failing to find any difference (Barlow, Rapee, & Brown, 1992; Borkovec & Mathews, 1988), condition size varied only from 9 to 13 ($M = 10.50$). Although the third investigation (White, Keenan, & Brooks, 1992) averaged 27 clients per condition, this is the only GAD study to use group therapy of unknown relevance to the rest of the GAD therapy literature. In the present study, cell size was determined by a power analysis on the basis of between-groups effect sizes (average effect size is 0.37) from our prior component control investigation (Borkovec & Costello, 1993). Assuming $\alpha = .05$, 25 clients per condition ensured power at .82 for detecting between-group differences. Cell sizes in the present study thus provided for nearly sufficient power in completer analyses and for more than adequate power in intent-to-treat analyses, so lack of power is an unlikely explanation for the results. The lack of condition differences also cannot be attributed to any differential effects of credibility, expectancy for improvement, therapeutic relationship qualities, or quality of CT (in CT and CBT), which may have counteracted differential efficacy; the conditions were equivalent on measures of those constructs. Similarly, adherence to protocol was extremely high in all three conditions, and the CT administered in CT alone and in CBT was provided with equivalent levels of quality.

Other plausible explanations for lack of condition differences include the possibility that the clients as a group may not have actually improved because of the specific therapies but rather because of (a) nonspecific or common factors present equally in all three conditions or (b) effects of history, maturation, repeated testing, statistical regression, instrument drift, unknown differential selection effects negating the effectiveness of random assignment, or interactions of selection with the other factors (Campbell & Stanley, 1963). The absence of a nonspecific control condition or of a waiting-list–no-treatment condition precludes unambiguously ruling out these possibilities. Given the outcomes of past GAD research and the effects of the present conditions, however, we feel that these explanations are not likely. To make this argument, Table 3 also provides within-group effect sizes from eight previous comparisons of CBT with conditions controlling for nonspecific factors and four previous comparisons of CBT to no-treatment conditions. As mentioned in the results section, each of our three conditions largely equaled or surpassed the average effects of prior CBT groups and was superior to previous component conditions. Table 3 also indicates that each condition surpassed previous nonspecific and no-treatment control conditions. This evidence supports a tentative conclusion that each of our conditions was quite effective as well as specific in its effects.

Moreover, like prior CBT research findings, improvements in anxiety and depression were maintained in each condition for 2 years after therapy, the vast majority of clients no longer met GAD criteria, and sufficient reduction in symptoms occurred such that only a small percentage of clients (18.96%) had sought any further treatment during the follow-up period.

Assuming that all three conditions were specifically effective, as argued above, a further rival hypothesis for lack of differential efficacy is that the supportive listening element in the CT and SCD conditions contributed in an additive or interactive way to the efficacy of these techniques. On the other hand, the type of supportive listening used for this control feature has been previously found to be significantly inferior by itself to both CBT and applied relaxation in our prior study (Borkovec & Costello, 1993), and its use as control filler in the applied relaxation condition in that prior study did not increment its efficacy sufficiently to match the level of clinically significant change shown by the CBT condition at follow-up. Research manipulating the presence or absence of supportive listening when used with CBT components would be necessary to rule out unambiguously this alternative explanation, however. Although the use of supportive-listening elements in such studies as well as in the present study reduces the external

---

4 We also recalculated correlations between IIP–C octants and endstate scores using the endpoint data. Of the 32 correlations, 19 (5 at post, 7 at 6-month, 2 at 12-month, and 5 at 24-month follow-up) were significant ($ps < .05$) and negative for posttherapy IIP–C scores (compared with 5 significant correlations for pretherapy IIP–C). After Sime's Bonferroni correction, 15 of these posttherapy IIP–C correlations remained significant (5 involving postendstate correlations with Vindictive/Self-centered, Intrusive/Needy, Dominingering/Controlling, Overly accommodating, and Non-assertive; 7 involving 6-month-endstate correlations with Vindictive/Self-centered, Intrusive/Needy, Dominingering/Controlling, Cold/Distant, Nonassertive, Overly accommodating, and Self-sacrificing; 1 involving 12-month-endstate correlations with Nonassertive and 2 involving 24-month-endstate correlations with Vindictive/Self-centered and Cold/Distant). Partial correlations controlling for pretherapy anxiety scores resulted in similar results, except that Overly accommodating was no longer associated with posttherapy and 6-month follow-up outcome.
validity of results for CBT components, their use is critical for internal validity.

A final alternative explanation for the lack of differences resides in the possibility that the two component conditions were particularly effective in and of themselves, raising outcomes to the level of CBT. Tables 2 and 3 support this view. Although less time was spent in active treatment in the component conditions than in CBT because of their use of supportive listening, these two component conditions actually did provide more active therapy time (16 hr) than any previous component condition in the GAD therapy literature. Because GAD involves interactions among cognitive, somatic, and affective responses, it originally seemed logical that targeting more of these systems (as in CBT) would make more of an impact than targeting only one. An alternative hypothesis, given the outcomes of the present study, is that several pathways exist for changing the entire intrapersonal system and that targeting some response processes in therapy for a sufficiently long period of time might therefore affect all of the other processes involved in the maintenance of anxiety. If this were the case, evidence should exist in prior GAD research of a relationship between treatment time in component conditions and whether or not differences were found between a component and CBT. From CBT therapy investigations that included only GAD clients and one or more component conditions, average amount of time devoted to the component conditions was calculated in (a) those studies finding evidence of greater change for CBT compared with its component (Borkovec et al., 1987; Borkovec & Costello, 1993; Butler et al., 1991; Durham et al., 1994) and (b) those investigations failing to find a difference (Barlow et al., 1992; White et al., 1992). Studies finding a difference averaged 9.25 hr (SD = 3.40) of component therapy time, whereas studies demonstrating equivalence averaged 13.50 hr (SD = 1.73), a difference approaching significance, \( t(6) = 2.23, p < .07 \). Thus there is evidence that brief component treatment may be insufficient to create the same degree of change throughout all of the anxiety response systems that a combined treatment targeting all systems can produce, whereas lengthy behavioral or cognitive component treatment may generate eventual, widespread, indirect change throughout other response systems, similar to that provided directly by a combined CBT condition. Future experimental research directly manipulating short and long therapy time in component and CBT conditions, along with frequent assessment of change in each response system throughout the duration of therapy, would provide an opportunity to test this hypothesis.

No evidence was found that increasing the amount of therapy time to allow for thorough CT incremented the effectiveness of our CBT. Two sources of information lead to that conclusion. First, our CBT condition did not yield percentages of high-endstate functioning that surpassed those of prior investigations. Second, although CBT showed somewhat greater within-group effect sizes relative to prior CBT treatments at posttherapy and at 6-month follow-up, it was largely equivalent to previous CBT conditions at later follow-up periods (see Table 3).

Our clinical research program has now spent 16 years attempting to refine, develop, and evaluate behavioral and CT methods for treating GAD. Outcomes from the present study suggest that we need to look elsewhere for ways of incrementing the effectiveness of psychological treatment for this disorder. Given that GAD clients spend considerable time living in, and responding to thoughts and images of, a nonexistent future (Borkovec et al., in press), one possibility involves the further development of methods for helping clients to attend to and live more fully in the present moment, a goal that our past CBT approaches have attempted to accomplish (see Borkovec & Costello, 1993). More systematic ways of doing this are currently being explored by Roemer and Orsillo (in press) in their incorporation of aspects of mindfulness and acceptance therapy into CBT techniques for GAD. Other promising developments that have given rise to therapy applications include a therapeutic focus on the GAD client’s characteristic intolerance of uncertainty (Ladouceur, Dugas, et al., 2000) and on the potential role of erroneous positive and negative beliefs about worry (Wells, 1999).

For us, however, the correlational association of the IIP–C scales with posttherapy and 6-month follow-up outcome (and subsequent follow-up outcome in the endpoint analyses; see Footnote 4) suggests an additional possible direction. Interpersonal behavior may be a significant element in the interacting response systems involved in GAD. Specifically, worry and anxiety may develop and/or be maintained because of problems in one’s relationships with others and/or with failures in having one’s interpersonally mediated needs met. Being domineering and vindictive (as in Horowitz et al.’s 1993 study) or intrusive in one’s relationships may be particularly associated with maladaptive emotional life. In a similar vein, Crits-Christoph, Connolly, Azarian, Crits-Christoph, and Shappell (1996) have offered persuasive arguments and open-trial data to support the possible efficacy of an interpersonally oriented psychodynamic therapy for GAD, and Durham, Allan, and Hackett (1997) have found evidence that quality of intimate relationships may predict long-term outcome in the treatment of GAD. The present study also indicates that interpersonal focus of CBT may not be ideal for affecting such interpersonal functioning. Although the clients showed improvement on the eight IIP–C octants from pretherapy to posttherapy, very few within-condition effects were significant after correction for Type I error. There thus may be potential therapeutic value in adding some form of interpersonal therapy to the CBT package. An outcome investigation is currently contrasting CBT with and without interpersonal therapy (see Newman, Castonguay, Borkovec, & Molnar, in press). Should the combined CBT and interpersonal therapy be found to be superior to CBT alone, evidence would exist for a causative role for interpersonal functioning in GAD and for a causative role for the therapeutic modification of that functioning.

References


Received March 2, 2001
Revision received June 5, 2001
Accepted July 30, 2001

---

**ORDER FORM**

Start my 2002 subscription to *Journal of Consulting and Clinical Psychology!* ISSN: 0022-006X

- $94.00, APA MEMBER/AFFILIATE
- $189.00, INDIVIDUAL NONMEMBER
- $434.00, INSTITUTION
  In DC add 5.75% / In MD add 5% sales tax

**TOTAL AMOUNT ENCLOSED** $

Subscription orders must be prepaid. (Subscriptions are on a calendar basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.

**SEND THIS ORDER FORM TO:**
American Psychological Association
Subscriptions
750 First Street, NE
Washington, DC 20002-4242

Or call (800) 374-2721, fax (202) 336-5568.
TDD/TTY (202) 336-6123.
Email: subscriptions@apa.org

- [ ] Send me a FREE Sample Issue
- [ ] Check Enclosed (make payable to APA)
- [ ] VISA  [ ] MasterCard  [ ] American Express

Cardholder Name ___________________________ Exp. date _____________

Card No. ___________________________ Signature (Required for Charge)

**BILLING ADDRESS:**

City ___________________________ State _____ Zip ___________

Daytime Phone ___________________________

**SHIP TO:**

Name ___________________________ Address ___________________________

City ___________________________ State _____ Zip ___________

APA Member # ___________________________ CCP412