

Effectiveness of Cognitive–Behavioral Treatment for Panic Disorder Versus Treatment as Usual in a Managed Care Setting: 2-Year Follow-Up

Michael E. Addis, Christina Hatgis, Esteban Cardemil, Karen Jacob, Aaron D. Krasnow,
and Abigail Mansfield
Clark University

Eighty clients meeting criteria for panic disorder and receiving either panic control therapy (PCT; M. G. Craske, E. Meadows, & D. H. Barlow, 1994) or treatment as usual (TAU) in a managed care setting were assessed 1 and 2 years following acute treatment. PCT was provided by therapists with little or no previous exposure to cognitive–behavioral therapies. Analyses of the full intent-to-treat sample revealed no significant differences between the treatments across the follow-up period. However, when treatment completer status was added as a moderator, those receiving PCT showed lower levels of panic severity and phobic avoidance and a greater likelihood of achieving and maintaining clinically significant change. Benzodiazepine use during follow-up was associated with greater panic severity for those clients who received PCT, but no such relationship was found for TAU clients. Results are discussed in relation to the dissemination and effectiveness of PCT as well as evidence-based psychotherapies more generally.

Keywords: panic disorder, dissemination, empirically supported treatments, psychotherapy, effectiveness

There are now a number of empirically supported psychotherapy treatments (ESTs) for a range of adult psychiatric disorders (Chambless & Hollon, 1998; DeRubeis & Crits-Christoph, 1998; Westen & Morrison, 2001). Many ESTs offer both well-defined treatment techniques that are typically operationalized in treatment manuals and methods for objectively evaluating treatment outcomes (Addis, 1997; Wilson, 1996). Several of these treatments have been shown to produce superior outcomes compared with no-treatment, waitlist, and nonspecific control groups in controlled clinical trials (Chambless & Hollon, 1998; DeRubeis & Crits-Christoph, 1998).

Although the efficacy of many ESTs continues to be evaluated, a distinction has emerged in the literature between treatment efficacy versus effectiveness (Addis, 2002; Carroll & Nuro, 2002; Street, Niederhe, & Lebowitz, 2000). Whereas efficacy refers to treatment outcomes obtained under tightly controlled research conditions, effectiveness describes treatment outcomes obtained under routine clinical practice conditions. Some researchers and practitioners have pointed to potentially critical differences between the contexts of controlled research and clinical practice that may limit the effectiveness of ESTs (e.g., Addis & Krasnow, 2000; Fensterheim & Raw, 1996; Persons & Silberschatz, 1998; Silverman, 1996). Accordingly, the field has seen a dramatic increase in the number of effectiveness studies for ESTs in both mental health and primary care settings (Addis et al., 2004; Franklin, Abramow-

itz, Kozak, Levitt, & Foa, 2000; Hahlweg, Fiegenbaum, Frank, Schroeder, & von Witzleben, 2001; Lincoln et al., 2003; Merrill, Tolbert, & Wade, 2003; Persons, Bostrom, & Bertagnolli, 1999; Tuschen-Caffier, Pook, & Frank, 2001; Wade, Treat, & Stuart, 1998; Warren & Thomas, 2001; Wells et al., 2004). A smaller percentage of these studies have compared outcomes of an EST conducted in a clinical practice setting with treatment as usual (TAU) in the same setting (Addis et al., 2004; Byford et al., 2003; Henggeler et al., 1999; Morgenstern, Blanchard, Morgan, Labouvie, & Hayaki, 2001; Sheidow et al., 2004; Wells et al., 2004). Because a major goal of evidence-based practice is to improve outcomes in real-world clinical settings, such comparisons are true litmus tests of the effectiveness and clinical utility of ESTs (Chorpita et al., 2002; Weisz & Addis, in press; Weisz, Southam-Gerow, Gordis, & Connor-Smith, 2003).

The purpose of the current study was to evaluate the long-term outcomes of clients participating in a randomized effectiveness trial comparing a well-established EST with routine clinical care. Addis et al. (2004) compared an empirically supported cognitive–behavioral treatment (CBT) for panic disorder (panic control therapy [PCT]; Craske, Meadows, & Barlow, 1994) with TAU in a capitated managed care clinical practice setting. Clients meeting criteria for panic disorder with a wide range of comorbid conditions were randomly assigned to a therapist recently trained in PCT or to a therapist conducting TAU. Medication use was not controlled in either condition. In both the intent-to-treat and the treatment completer samples, clients receiving treatment from a therapist trained in PCT showed greater change from pre- to posttreatment on both panic-related symptoms and general well-being. Among treatment completers, an average of 42% of those in PCT and 18.8% in TAU achieved clinically significant change (Jacobson & Truax, 1991). In the current study, we hypothesized that at 1- and 2-year follow-ups, clients who initially received PCT would show greater symptom reduction and better functioning than clients who received TAU.

Michael E. Addis, Christina Hatgis, Esteban Cardemil, Karen Jacob, Aaron D. Krasnow, and Abigail Mansfield, Department of Psychology, Clark University.

Preparation of this article was supported by National Institute of Mental Health Grant R29 MH57778.

Correspondence concerning this article should be addressed to Michael E. Addis, Department of Psychology, Clark University, 950 Main Street, Worcester, MA 01610. E-mail: maddis@clarku.edu

Method

Participants

Details of participant and therapist recruitment were reported in full in Addis et al. (2004). All participants were members of a large HMO that serves a wide cross-section of central Massachusetts. Eligible participants met diagnostic criteria for panic disorder with or without agoraphobia. Although all participants entered into the study met full criteria for panic disorder, those who were subthreshold for a strict diagnosis of panic disorder but identified panic symptoms as their primary reason for seeking treatment would have also been eligible for the study. No exclusions were made on the basis of medication use for anxiety or for other comorbid psychological or medical problems. The sample was composed of 80 participants (70% women; mean age = 39.9 years, $SD = 12.9$, range = 18–70). Thirty-eight participants were randomly assigned to PCT, and 42 were assigned to TAU.

Therapists

Ten therapists participated in the study, the majority of whom were master's-level practitioners with little or no experience in CBT treatment of anxiety or depression. None of the 10 therapists identified their primary theoretical orientation as cognitive-behavioral; as a group, they were approximately equally distributed among eclectic, family systems, psychodynamic, and humanistic orientations in their self-descriptions. We created two groups of five therapists, matching each according to number of years in clinical practice and prior exposure to CBT of anxiety or depression (e.g., attending a workshop). By random assignment, one group was appointed to conduct PCT and the other to conduct TAU. Those in the TAU condition were told that they would receive training in PCT upon completion of the acute treatment phase of the study. Prior to agreeing to participate in the study, all therapists were informed of the processes for assigning treatment conditions (for both therapists and clients). The level of therapist training provided was consistent with the goals of an effectiveness study. PCT therapists received a 2-day workshop followed by audiotape review and supervision of two training cases by an expert in PCT who was not affiliated with the study. Subsequent optional biweekly group supervision was provided by the principal investigator. Further information on therapist recruitment and the details and rationale for the training protocol were presented in Addis et al. (2004).

Treatments

PCT (Craske et al., 1994) is a manual-guided 12- to 15-session CBT protocol. The treatment includes education about the causes and maintenance of panic disorder, breathing retraining, and cognitive restructuring, as well as interoceptive and agoraphobic exposure components. PCT therapists were encouraged but not compelled to use the treatment protocol. We explained to therapists that the treatment typically requires 12 to 15 sessions and is designed to be conducted on a weekly basis. Therapists were not, however, required to schedule a specific number or sequence of treatment sessions. Treatment was paid for through the client's existing mental health benefits with the exception of a \$10 copayment for each session that was covered by the study.

TAU therapists were instructed to provide whatever treatment they deemed appropriate for their clients. In both the TAU and PCT conditions, decisions about medication use were left up to clients, their therapists, and other medical or psychiatric providers involved in their cases. During treatment, medication referrals generally were made to psychiatrists working within the clinic. Some clients, however, may also have received medication from primary care or other physicians.

Procedure

All participants were assessed 1 and 2 years following their designated posttreatment date, which occurred 5.5 months following assignment to

treatment condition. Participants completed a battery of self-report measures and participated in two structured interviews (described below) conducted by doctoral-level interviewers blind to the client's treatment condition. All participants were compensated with \$35 for the 1-year follow-up and with \$60 for the 2-year.

Measures

Panic severity. The Panic Disorder Severity Scale (PDSS; Shear et al., 2001) is a 7-item semistructured interview assessing overall severity of panic disorder and agoraphobic avoidance. Information is also gathered about severity of panic attacks, anticipatory anxiety, avoidance behavior, and work and social dysfunction. Psychometric evaluations indicate that the PDSS has good interrater reliability, moderate internal consistency, and good construct validity (Shear et al., 2001). The PDSS was completed by a trained graduate-level clinical assessor at the 1- and 2-year follow-ups.

Phobic avoidance. The Fear Questionnaire (FQ; Marks & Matthews, 1979) is a 24-item self-report questionnaire that assesses phobic avoidance on three anxiety subscales (Agoraphobia, Blood/Injury, Social Phobia) and one general Anxiety and Depression subscale. We used the total phobia score, which combines items from the three anxiety subscales, for all analyses. Psychometric properties for the FQ have been reported as adequate (Cox, Parker, & Swinson, 1996), and the measure is recommended for studies assessing phobic avoidance in panic disorder (Cox, Swinson, Parker, Kuch, & Reichman, 1993; Shear & Maser, 1994). Participants completed the FQ at both the 1- and 2-year follow-ups.

Depression. The Beck Depression Inventory (BDI; Beck & Steer, 1987), a 21-item self-report questionnaire, assesses the severity of depressive symptoms over a 2-week period. This measure is widely used; is highly correlated with other measures of depression, such as the Hamilton Rating Scale (Hamilton, 1967); and has excellent psychometric properties (Beck, Steer, & Garbin, 1988). Participants completed the BDI at both the 1- and 2-year follow-ups.

General well-being. The Outcome Questionnaire (OQ-45; American Professional Credentialing Services, 1996) is a 45-item self-report measure assessing symptom distress, social role functioning, and interpersonal functioning. We chose the OQ-45 as a measure of general well-being to complement the other more panic-specific measures. The OQ-45 demonstrates good internal consistency, test-retest reliability, and concurrent validity with other measures of psychiatric symptoms and interpersonal functioning (American Professional Credentialing Services, 1996; Mueller, Lambert, & Burlingame, 1998; Umphress, Lambert, Smart, Barlow, & Clouse, 1997). In all analyses, we used the total score combining the three subscales. Participants completed the OQ-45 at both the 1- and 2-year follow-ups.

Psychotherapy and medication use interview. This structured interview was constructed for the current study to serve as a measure of psychotherapy sessions attended and psychotropic medication used during the follow-up period. Doctoral student interviewers administered the interview at the 1- and 2-year follow-ups. Medication and psychotherapy use was assessed monthly in a follow-back format beginning with the month previous to the assessment. For each month, interviewers assessed and recorded whether clients were taking any psychotropic medications and, if so, the types of medications and the frequency (less than once a week, a few times per week, nearly every day, once per day, bid, tid, qid, and more than four times per day). Because many clients had difficulty determining the exact dosage of medications, the measure of frequency was taken as more reliable. Outcomes are presented for selective serotonin reuptake inhibitors (SSRIs) and benzodiazepines, by far the two most commonly used medications. Over 90% of clients taking SSRIs took them once per day. We therefore collapsed the scale for SSRI use into a binary code, with a one indicating that clients were taking the medication during the month and a zero indicating nonuse. For psychotherapy use, interviewers assessed the type of therapy received (e.g., psychotherapy for panic disorder, psy-

chotherapy for depression, or marital therapy) during each month of the follow-up period and the number of sessions attended for each treatment.

Results

Client Subsamples

The full intent-to-treat sample included all participants enrolled into the study ($N = 80$), regardless of how many sessions they attended. We also considered treatment completer status as a moderator variable in the primary analyses. Eleven participants (5 PCT, 6 TAU) did not attend any treatment sessions. All of these participants completed posttreatment assessments, and their data were included in the intent-to-treat analyses. Analyses were also completed for subsamples of "treatment completers" who attended at least eight sessions ($M = 10.8$, $SD = 2.6$) by 8.5 months following assignment to treatment condition. Because length of treatment in TAU was expected to vary according to clients' and therapists' judgment, the designation of eight sessions as a cutoff was based primarily on the need to equate the two treatment conditions for analytic purposes. In addition, if therapists covered an average of 1.5 chapters of the client workbook per week, eight sessions would have provided sufficient time to make it through the three critical components of PCT, including interoceptive exposure (Craske et al., 1994).

Missing Data

We were unable to locate 5 clients in TAU who consequently did not complete the 1-year follow-up. All clients in PCT completed the 1-year follow-up, although 1 client did not complete the OQ-45 and 2 did not complete the PDSS (these 2 clients refused to be contacted by phone but completed the self-report measures and returned them in the mail). At the 2-year follow-up, 6 clients in TAU were unable to be contacted and did not complete the follow-up assessment (5 of these were the same clients who were missing at the 1-year follow-up). In PCT, 2 clients did not complete the full follow-up assessment. In addition, 2 clients did not complete the OQ-45. Clients with any missing data points over the 2-year follow-up ($n = 11$) had higher posttest PDSS scores ($M = 10.6$, $SD = 8.7$) than did clients without missing data ($M = 5.2$, $SD = 4.7$), $F(1, 77) = 8.9$, $p < .01$. No other differences were found between clients with and without missing data during the follow-up period.

Statistical Procedures

Means and standard deviations for scores on each of the primary outcome measures at the 1- and 2-year follow-ups are presented in Table 1. We used hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) to analyze change in symptoms over the course of the study. HLM is ideal for analyzing change over time because it accommodates missing data among repeated measurements using empirical Bayesian estimates. HLM also makes fewer unrealistic assumptions regarding within-subject correlations and change in correlations over time (Raudenbush & Bryk, 2002). All analyses were conducted by using the Proc Mixed module of SAS (e.g., Singer, 1998). We focused primarily on panic severity as measured by the PDSS but also replicated the analyses for other dependent measures.

Table 1
Means and Standard Deviations for Outcome Measures in Both the Intent-to-Treat and Treatment Completer Samples

Measure	PCT			TAU			d^a
	n	M	SD	n	M	SD	
Intent-to-treat sample ($N = 80$)							
PDSS							
1 year	38	6.8	5.6	42	6.0	6.0	.14
2 years	38	5.6	6.2	42	6.3	6.8	.11
FQ							
1 year	38	29.2	20.8	39	30.9	21.6	.08
2 years	38	31.2	23.9	39	29.7	20.6	.07
BDI							
1 year	38	11.0	8.5	38	7.9	7.4	.39
2 years	38	9.3	8.4	38	9.2	8.4	.01
OQ-45							
1 year	37	62.5	23.7	39	56.5	23.6	.25
2 years	37	56.0	29.0	39	56.1	24.6	.00
Treatment completers ($n = 32$)							
PDSS							
1 year	20	5.2	5.1	12	7.2	6.0	.36
2 years	20	3.9	4.7	12	7.3	7.2	.58
FQ							
1 year	20	22.8	20.6	11	33.4	16.7	.54
2 years	20	22.7	20.2	11	33.1	17.6	.53
BDI							
1 year	20	9.5	8.0	11	6.7	4.8	.39
2 years	20	8.0	8.9	11	6.6	6.4	.18
OQ-45							
1 year	20	58.5	24.0	11	51.0	15.0	.35
2 years	20	53.5	31.9	11	47.8	16.5	.21

Note. PCT = panic control therapy; TAU = treatment as usual; PDSS = Panic Disorder Severity Scale; FQ = Fear Questionnaire Total Phobia Scale; BDI = Beck Depression Inventory; OQ-45 = Outcome Questionnaire.

^a Refers to between-treatment effect sizes calculated according to Cohen (1977).

Our analyses in the current study directly extend the results reported in the original acute treatment outcome analyses. Intent-to-treat analyses of the posttreatment effects revealed no differences between PCT and TAU. However, a differential treatment effect emerged when treatment completer status was taken into consideration (Addis et al., 2004); clients who received PCT and attended at least eight sessions reported significantly lower PDSS scores than did clients receiving TAU. Thus, in the HLM analyses reported in this article, we continued to examine the extent to which treatment completer status interacted with treatment condition to predict outcome over the follow-up period. In addition, because we were primarily interested in examining the course of symptoms over the follow-up period, we analyzed time in a piecewise manner in which we conceptually divided the study into the acute phase and the follow-up phase. This analytic strategy allowed us to examine symptom change over time separately for the acute treatment period and the follow-up period (see Willett, Singer, & Martin, 1998).

We therefore began our analyses with a model that included two Level 1 time variables (acute treatment time and follow-up time)

and a single Level 2 variable (condition assignment). In this model, change in PDSS scores over the acute period can be understood along three parameters: (a) linear change for those individuals who received TAU (β_{10}), (b) linear change for individuals who received PCT (β_{11}), and (c) unexplained error. Change in PDSS scores over the follow-up period can be understood along three similar parameters: (a) linear change for TAU (β_{20}), (b) linear change for PCT (β_{21}), and (c) unexplained error.

We then expanded the model to add treatment completer status and its interaction with condition assignment as additional Level 2 variables. In the full model, change in PDSS scores over the acute period can be understood along five parameters: (a) linear change for those individuals who received TAU and were not treatment completers (β_{10}), (b) linear change for individuals who received PCT and were not treatment completers (β_{11}), (c) linear change for individuals who received TAU and were treatment completers (β_{12}), (d) linear change for individuals who received PCT and were treatment completers (β_{13}), and (e) unexplained error. Change in PDSS scores over the follow-up period can be understood along five similar parameters: (a) linear change for TAU noncompleters (β_{20}), (b) linear change for PCT noncompleters (β_{21}), (c) linear change for TAU completers (β_{22}), (d) linear change for PCT completers (β_{23}), and (e) unexplained error.

Panic Symptoms

We began by examining change in the complete intent-to-treat sample and therefore did not include treatment completer status in the equations. This analysis revealed no difference in rate of change in PDSS scores between TAU and PCT over the acute period ($\beta_{11} = -0.23$, $t(302) = -1.29$, *ns*), or the follow-up period ($\beta_{21} = -0.01$, $t(302) = -0.10$, *ns*). When we expanded the analyses to the full model, however, there was a significant difference in rate of change in PDSS scores during the acute period ($\beta_{13} = -0.8289$, $t(298) = -2.11$, $p < .05$). Probing these analyses indicated that individuals in PCT who were completers were reporting the most rapid rate of change of the four subgroups. These results using HLM replicated the repeated measures analyses reported by Addis et al. (2004).

Examining change over the 2-year follow-up period (controlling for change over the acute phase), we found that none of the four subgroups (i.e., TAU noncompleters, TAU completers, PCT noncompleters, and PCT completers) demonstrated statistically significant change from posttreatment through the follow-up period (β_{20} , β_{21} , β_{22} , and β_{23} were all nonsignificant), indicating that the differences between the four subgroups during the acute phase were maintained through the follow-up. Figure 1 provides a visual representation of these results. We next conducted two separate

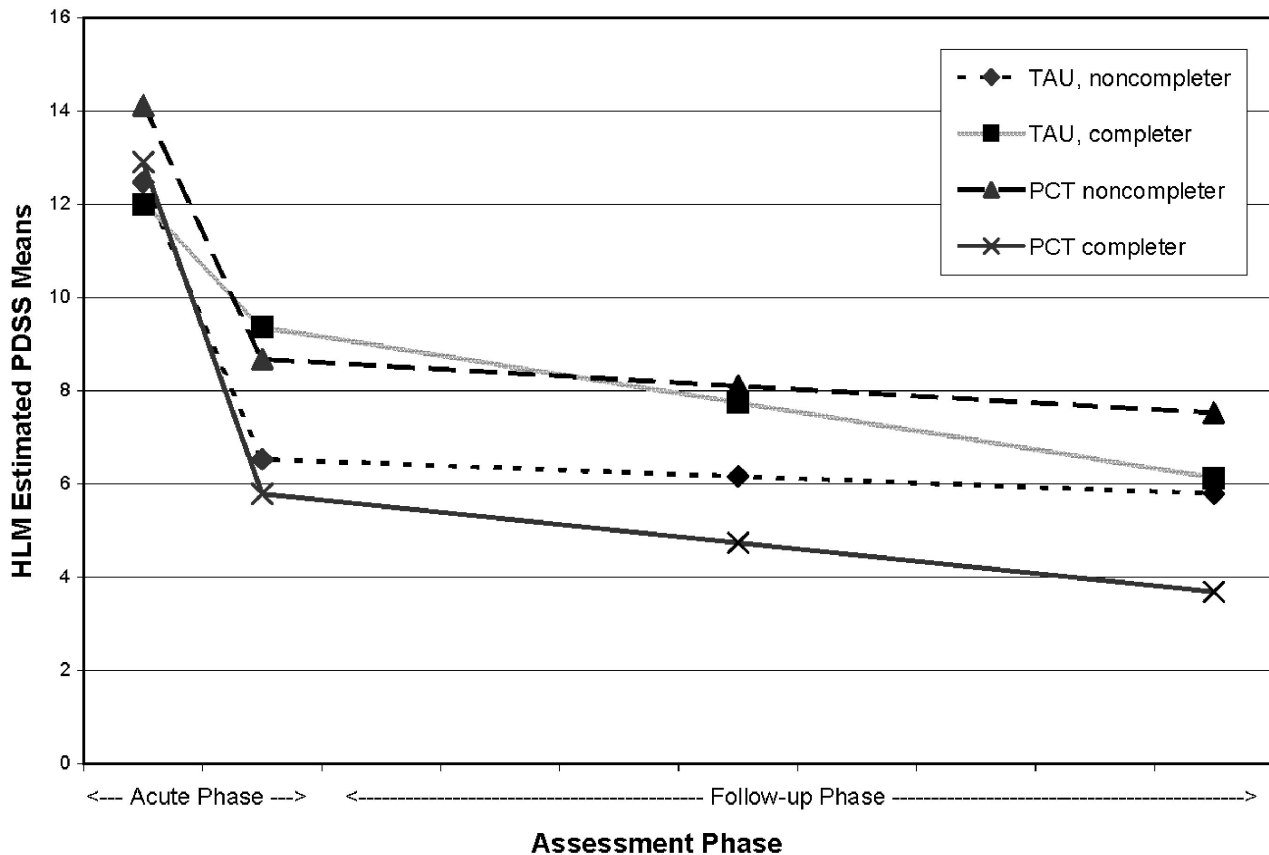


Figure 1. Simple means and simple slopes for Panic Disorder Severity Scale (PDSS) scores in four treatment subsamples. HLM = hierarchical linear modeling; TAU = treatment as usual; PCT = panic control therapy.

intercept analyses in which we first set the intercept to the 1-year follow-up period and then set it to the 2-year follow-up period. These analyses allowed us to compare PDSS scores among the four subgroups at those specific time points. At the 1-year follow-up period, there were statistically significant differences among the four conditions ($\beta_{03} = -4.95$), $t(76) = -2.10$, $p < .05$. The pattern was the same at the 2-year follow-up, although the difference was just short of statistical significance ($\beta_{03} = -4.1775$), $t(76) = -1.40$, $p = .16$. Probing these analyses indicated that in each case, PCT completers were reporting the lowest PDSS scores among the four subgroups.

The best-fitting model for these analyses was one in which we allowed the intercept, acute time, and follow-up time to vary randomly: when compared with the null model, $\chi^2(6, N = 80) = 186.09$, $p < .0001$; when compared with allowing only acute time to vary randomly, $\chi^2(3, N = 80) = 17.70$, $p < .001$; when compared with allowing only follow-up time to vary randomly, $\chi^2(3, N = 80) = 8.40$, $p < .05$. In this model, there was statistically significant random variation for the intercept ($\tau_{00} = 25.19$, $Z = 5.03$, $p < .0001$), acute time ($\tau_{11} = 0.27$, $Z = 2.13$, $p < .05$), and follow-up time ($\tau_{22} = 0.04$, $Z = 3.14$, $p < .001$). These random factors suggested that there remained significant variation in both panic symptoms and change in panic symptoms over time, after accounting for the predictor variables.

Additional Outcome Measures

We conducted the same HLM analyses for change in BDI scores, the Phobia subscale of the FQ, and the OQ-45. For both BDI scores and OQ-45 scores, treatment condition and treatment completer status were unrelated to linear change during either the acute treatment phase or the 2-year follow-up. In addition, treatment condition and treatment completer status were unrelated to final outcome at the 2-year follow-up period. For scores on the Phobia subscale, there was also no treatment condition or treatment completer effect on linear change during either the acute treatment phase or the 2-year follow-up. However, intercept analyses indicated significant differences in scores on the Phobic subscale at both the 1-year follow-up point ($\beta_{03} = -20.45$), $t(76) = -2.00$, $p < .05$, and the 2-year follow-up point ($\beta_{03} = -22.62$), $t(76) = -2.20$, $p < .05$. In both cases, probing these interactions indicated that PCT completers were reporting lower scores than the other three subgroups.

Medication and Psychotherapy Use

We found no significant differences between PCT and TAU in the degree to which clients in either treatment used SSRIs, benzodiazepines, or additional psychotherapy sessions during the follow-up period. To investigate the extent to which medication use might influence the panic severity during the follow-up, we conducted exploratory HLM analyses with benzodiazepine use and SSRI use as time-varying covariates. Because we had no clear prediction about how medication use would interact with completer status, and to reduce the number of exploratory analyses, we did not include treatment completer status in this model. In addition, because we collected data on medication use only during the follow-up period, we did not conduct a piecewise analysis, instead using only one time variable that represented linear change from

posttreatment through the 2-year follow-up. We included benzodiazepine use, SSRI use, and psychotherapy use as additional time-varying Level 1 variables. Treatment condition was the single Level 2 variable. We found that neither SSRI use nor psychotherapy use predicted change in PDSS scores over the 2-year follow-up, above and beyond the effect of time. There was a significant difference, however, with benzodiazepine use ($\beta_{31} = 0.81$), $t(139) = 2.03$, $p < .05$. Probing these analyses revealed that for participants in the PCT condition, benzodiazepine use was associated with significantly higher PDSS scores (i.e., worsening of symptoms). In this case, the best-fitting model was one in which the intercept, SSRI use, and benzodiazepine use varied randomly: when compared with the null model, $\chi^2(6, N = 80) = 83.24$, $p < .0001$. There was statistically significant random variation for the intercept ($\tau_{00} = 24.23$, $Z = 4.14$, $p < .0001$) and SSRI use ($\tau_{11} = 18.49$, $Z = 2.01$, $p < .05$), but not for benzodiazepine use, suggesting that significant variation remained in panic symptoms after accounting for the predictor variables.

Clinical Significance Analyses

Following the methods described by Jacobson and Truax (1991), we calculated the percentage of clients in each treatment condition that achieved reliable change and end-state functioning within a nonclinical distribution at posttreatment and that maintained those gains at the 1- and 2-year follow-ups. Cut scores from published norms were obtained for the PDSS (Shear et al., 2001), the OQ-45 (American Professional Credentialing Services, 1996), the FQ (Gillis, Haaga, & Ford, 1995), and the BDI (Seggar, Lambert, & Hansen, 2002). We calculated reliable change scores based on the internal consistency or test-retest reliability of each measure.

The clinical significance analyses were designed to answer the question, What percentage of clients in each treatment condition improved significantly by posttreatment and remained improved during the follow-up period? For clients to be counted as demonstrating clinically significant change at the 1-year follow-up, they had to have met the cutoff criteria at both posttest and 1 year.¹ For clients to be counted as demonstrating clinically significant change at the 2-year follow-up, they had to have met the cutoff criteria at posttest, 1 year, and 2 years. Thus, the percentages presented in Table 2 represent the number of clients initially entering treatment who achieved clinically significant improvement at posttest and maintained it at a particular follow-up point.

Table 2 presents the results of clinical significance analyses for both the intent-to-treat sample and the sample of treatment completers. Within the intent-to-treat sample, 39.5% of clients in PCT and 23.8% of clients in TAU maintained clinically significant improvement from posttreatment to the 1-year follow-up as measured by the PDSS (Fisher's exact $p = .10$). Rates of clinically significant improvement were considerably lower on the other three primary outcomes measures, averaging 8.5% with no significant differences in rates of improvement between the two treatments. At the 2-year follow-up, 36.8% of clients in PCT and

¹ At 1 and 2 years, clinically significant change was calculated by comparing change from pretreatment with the follow-up period on a particular measure.

Table 2
Percentage of Participants Maintaining Clinically Significant Change at Each Follow-Up Period

Period	PCT		TAU		Fisher's exact <i>p</i>
	<i>n</i>	%	<i>n</i>	%	
Intent-to-treat sample (<i>N</i> = 80)					
1 year					
PDSS	15	39.5	10	23.8	.10
FQ	5	13.2	2	4.8	.18
BDI	3	7.9	2	4.8	.45
OQ-45	5	13.2	3	7.1	.30
2 years					
PDSS	14	36.8	9	21.4	.10
FQ	2	5.3	1	2.4	.46
BDI	1	2.6	1	2.4	.73
OQ-45	4	10.5	3	7.1	.44
Treatment completers (<i>n</i> = 32)					
1 year					
PDSS	11	55.0	1	8.3	<.01
FQ	4	20.0	0	0	.14
BDI	1	5.0	0	0	.63
OQ-45	4	20.0	0	0	.14
2 years					
PDSS	11	55.0	1	8.3	<.01
FQ	1	5.0	0	0	.63
BDI	1	5.0	0	0	.63
OQ-45	4	20.0	0	0	.14

Note. PCT = panic control therapy; TAU = treatment as usual; PDSS = Panic Disorder Severity Scale, 5.5- or 8.5-month follow-up; FQ = Fear Questionnaire Total Phobia Scale; BDI = Beck Depression Inventory; OQ-45 = Outcome Questionnaire.

21.4% of clients in TAU had maintained clinically significant improvement since posttest (Fisher's exact $p = .10$). As with the 1-year follow-up results, rates of maintained clinically significant improvement from posttreatment to the 2-year follow-up were noticeably lower on all other measures, averaging 5.1% with no significant differences between the treatments.

Within the treatment completer sample, 55.0% of clients in PCT and 8.3% in TAU achieved clinically significant improvement in panic severity by posttreatment and maintained it at the 1-year follow-up (Fisher's exact $p > .01$). These rates did not change by the 2-year follow-up. Clients in PCT also appeared to show a greater rate of stable clinically significant improvement on the OQ-45 (20%) compared with clients in TAU (0%), although this difference was not statistically significant (Fisher's exact $p = .14$).

Finally, we compared rates of clinically significant improvement over the course of the follow-up in PCT completers versus the rest of the sample (PCT noncompleters and all clients receiving TAU). This analysis tested whether those clients receiving the highest dose of PCT achieved superior outcomes compared with all other clients receiving treatment. Completers in PCT showed significantly greater rates of clinically significant improvement on the PDSS at the 1-year (55% vs. 23.3%, Fisher's exact $p < .01$) and 2-year follow-ups (55% vs. 20%, Fisher's exact $p < .01$). On the OQ-45, 20% of PCT completers and 5% of TAU completers demonstrated clinically significant change (Fisher's exact $p =$

.06). No other significant differences in rates of clinically significant improvement emerged when comparing PCT completers with the rest of the sample.

Discussion

The purpose of this study was to assess the long-term outcomes of clients receiving either PCT (Craske et al., 1994) or TAU in a managed care setting. Acute treatment phase outcomes supported the effectiveness of PCT in clinical practice and suggested a modest advantage of PCT over usual care (Addis et al., 2004). Results from the current study indicate, in general, that the positive effects of PCT are maintained through the 2-year follow-up period. HLM analyses revealed no differences between the treatment conditions in the rate of change during the follow-up period in either the intent-to-treat or the treatment completer samples. However, treatment completers in PCT showed significantly lower scores on panic severity and phobic avoidance at the 1-year follow-up and significantly lower scores on phobic avoidance at the 2-year follow-up.

We had hypothesized that clients receiving PCT would show greater change from posttreatment through the follow-up period compared with those receiving TAU. The results revealed no differences between the treatments in degree of change from posttreatment to the 1- or 2-year follow-ups in any of the client subsamples. In fact, there were no significant main effects for time in any of the analyses, indicating that scores on each of the outcome measures remained relatively stable from posttreatment through the follow-up. These results, combined with the acute treatment phase outcomes, suggest that the majority of change clients achieved occurred during the initial phase of treatment; we found no evidence of delayed treatment effects in either PCT or TAU. Nor did we find any evidence of deterioration on average in either treatment condition. Thus, the current results suggest that the bulk of change achieved by clients receiving treatment for panic disorder in clinical practice settings is achieved during the initial treatment episode (e.g., 5.5–8.5 months in the current study). Because the current study was conducted in a single practice setting with a relatively small ($N = 10$) number of therapists, replication is necessary before firm conclusions can be drawn about typical patterns of treatment gains in clinical practice among clients with panic disorder.

Although we found no evidence of differential change over the follow-up, differences between the treatments did emerge when comparing mean scores on outcome measures at a particular follow-up period. Within treatment completers, those receiving PCT showed lower levels of panic severity at the 1-year follow-up. In addition, treatment completers receiving PCT also reported lower levels of phobic avoidance at the 1- and 2-year follow-ups as indicated by self-report scores on the FQ. No differences were found on measures of depression or general well-being. In addition, no differences between the treatments were found on any measure within the intent-to-treat sample, suggesting that the differential effectiveness of PCT was limited to those clients who attended at least eight sessions during the acute treatment phase. These results are consistent with the acute treatment phase outcomes (Addis et al., 2004) and suggest that among those clients attending at least eight sessions of treatment, PCT produces superior outcomes that are maintained across a 2-year follow-up. It is

worth noting that for practical reasons our definition of treatment completer (i.e., attending at least eight sessions) fell short of what is typically recommended for PCT (Craske et al., 1994). Given a documented dose–response relationship in ESTs, it is possible, and perhaps likely, that a stronger effect would have emerged if a greater percentage of participants had received the full PCT package (Hansen, Lambert, & Forman, 2002).

We used a stringent definition of clinically significant change during the follow-up period; clients needed to have achieved clinically significant change by posttreatment and maintained it at a particular follow-up period. Within the intent-to-treat sample, 40% of clients receiving PCT and 24% receiving TAU demonstrated clinically significant change in panic severity from pretreatment to the 1-year follow-up. By the 2-year follow-up, the rates were 37% for PCT and 21% for TAU. Thus, these results suggest that within the intent-to-treat sample, approximately one third of clients in PCT and one fifth of clients in TAU were able to achieve clinically significant change and maintain it across a 2-year follow-up. Among treatment completers, clinically significant changes in panic severity were considerably more common in PCT (55%) than in TAU (8%) at both the 1- and 2-year follow-ups. Within this sample, 70% of clients in PCT and 41% in TAU had achieved clinically significant change in panic severity 8.5 months after assignment to treatment condition (Addis et al., 2004). These results are consistent with a greater long-term prophylactic effect for PCT following treatment of at least eight sessions; treatment completers in PCT achieved higher rates of clinically significant improvement from pretest to the 2-year follow-up than did any other client subsample in either PCT or TAU.

We found no evidence of differences between the treatment conditions in the degree of client change from posttreatment to either follow-up period. These results, taken together with the mean differences between PCT and TAU at posttreatment and follow-up, indicate that treatment gains achieved by posttest were relatively stable and continued to be greater in PCT than in TAU, at least among treatment completers. These results are encouraging and suggest that a minimum of at least eight sessions of PCT during an acute treatment phase may, on average, produce a prophylactic effect over the longer term. The rates of clinically significant improvement from pretreatment throughout the 2-year follow-up among treatment completers are consistent with this interpretation.

We also examined the associations between additional treatment and clinical status during the follow-up period. In general, we found no linkages between medication or psychotherapy use and symptom severity on any measure during the follow-up period. The one exception was that benzodiazepine use during follow-up was associated with increased panic severity for clients who received PCT, whereas no such relationship was found for clients who had received TAU. These findings echo other studies demonstrating either no benefit or negative effects of combining benzodiazepine use with CBT for panic (Brown & Barlow, 1995; Otto, Pollack, & Sabatino, 1996; Westra, Stewart, & Conrad, 2002). They suggest further that the negative effects of combining CBT and benzodiazepines may continue after treatment has ended. It should be noted that in the current effectiveness study, it was unclear both how the combination of CBT and medication use was framed for different clients and how the treatments were sequenced. In contrast, there is evidence of efficacy for treatments

that combine benzodiazepine use and CBT sequentially with medication tapering as part of the treatment (Otto et al., 1993; Spiegel, Bruce, Gregg, & Nuzzarello, 1994).

There are some limitations to the current study that need to be considered when interpreting the results. First, differential attrition is always a concern when interpreting results of randomized trials, particularly with longer term outcomes. Our overall retention rate across the follow-up was fairly high (94% at 1 year, 90% at 2 years), which suggests that differential attrition was not a factor. Second, it is clear that many clients in both conditions sought additional treatment during the follow-up that may have affected their long-term outcomes. In general, we did not find associations between additional treatment seeking and clinical status during the follow-up, although the relationship may have been more complex than we were able to detect. Third, our estimates of treatment utilization during the follow-up period were based on self-report. The 12-month follow-back period at each assessment was fairly lengthy, and retrospective biases may have detracted from the accuracy of these reports. Greater accuracy may have been obtained by using actual medical and pharmaceutical records. This was our original intention, but it proved unfeasible because of a variety of difficulties related to working within the particular managed care system. Despite these limitations, the current results support the long-term effectiveness of PCT as delivered in a clinical service setting. They also suggest superior long-term outcomes for PCT compared with TAU among clients receiving at least eight sessions of treatment during an acute treatment phase.

References

- Addis, M. E. (1997). Evaluating the treatment manual as a means of disseminating empirically validated psychotherapies. *Clinical Psychology: Science and Practice*, 4, 1–11.
- Addis, M. E. (2002). Methods for disseminating research products and increasing evidence based practice: Promises, obstacles, and future directions. *Clinical Psychology: Science and Practice*, 9, 381–392.
- Addis, M. E., Hatgis, C., Krasnow, A. D., Jacob, K., Bourne, L., & Mansfield, A. (2004). Effectiveness of cognitive–behavioral treatment for panic disorder versus treatment as usual in a managed care setting. *Journal of Consulting and Clinical Psychology*, 72, 625–635.
- Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology*, 68, 331–339.
- American Professional Credentialing Services. (1996). *Administration and scoring manual for the OQ-45.2*. Washington, DC: Author.
- Beck, A. T., & Steer, R. A. (1987). *Manual for the revised Beck Depression Inventory*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8, 77–100.
- Brown, T. A., & Barlow, D. H. (1995). Long-term outcome in cognitive–behavioral treatment of panic disorder: Clinical predictors and alternative strategies for assessment. *Journal of Consulting and Clinical Psychology*, 63, 754–765.
- Byford, S., Knapp, M., Greenshields, J., Ukoumunne, O. C., Jones, V., & Thompson, S. (2003). Cost-effectiveness of brief cognitive behaviour therapy versus treatment as usual in recurrent deliberate self-harm: A decision-making approach. *Psychological Medicine*, 33, 977–986.
- Carroll, K. M., & Nuro, K. F. (2002). One size cannot fit all: A stage model for psychotherapy manual development. *Clinical Psychology: Science and Practice*, 9, 396–406.

- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*, 7–15.
- Chorpita, B. F., Yim, L. M., Donkervoet, J. C., Arensdorf, A., Amundsen, M. J., McGee, C., et al. (2002). Toward large-scale implementation of empirically supported treatments for children: A review and observations by the Hawaii Empirical Basis to Services Task Force. *Clinical Psychology: Science and Practice, 9*, 165–190.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.
- Cox, B. J., Parker, J. D. A., & Swinson, R. P. (1996). Confirmatory factor analysis of the Fear Questionnaire with social phobia patients. *British Journal of Psychiatry, 168*, 497–499.
- Cox, B. J., Swinson, R. P., Parker, J. D. A., Kuch, K., & Reichman, J. T. (1993). Confirmatory factor analysis of the Fear Questionnaire in panic disorder with agoraphobia. *Psychological Assessment, 5*, 235–237.
- Craske, M. G., Meadows, E., & Barlow, D. H. (1994). *Therapist's guide for the mastery of your anxiety and panic II & agoraphobia supplement*. Albany, NY: Graywind Publications.
- DeRubeis, R. J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology, 66*, 37–52.
- Fensterheim, H., & Raw, S. D. (1996). Psychotherapy research is not psychotherapy practice. *Clinical Psychology: Science and Practice, 3*, 168–171.
- Franklin, M. E., Abramowitz, J. S., Kozak, M. J., Levitt, J. T., & Foa, E. B. (2000). Effectiveness of exposure and ritual prevention for obsessive-compulsive disorder: Randomized compared with nonrandomized samples. *Journal of Consulting and Clinical Psychology, 68*, 594–602.
- Gillis, M. M., Haaga, D. A. F., & Ford, G. T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire, and Social Phobia and Anxiety Inventory. *Psychological Assessment, 7*, 450–455.
- Hahlweg, K., Fiegenbaum, W., Frank, M., Schroeder, B., & von Witzleben, I. (2001). Short- and long-term effectiveness of an empirically supported treatment for agoraphobia. *Journal of Consulting and Clinical Psychology, 69*, 375–382.
- Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology, 6*, 276–296.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implication for treatment delivery services. *Clinical Psychology: Science and Practice, 9*, 329–343.
- Henggeler, S. W., Rowland, M. D., Randall, J., Ward, D. M., Pickrel, S. G., Cunningham, P. B., et al. (1999). Home-based multisystemic therapy as an alternative to the hospitalization of youth in psychiatric crisis: Clinical outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 1331–1339.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- Lincoln, T. M., Rief, W., Hahlweg, K., Frank, M., von Witzleben, I., & Schroeder, B. (2003). Effectiveness of an empirically supported treatment for social phobia in the field. *Behaviour Research and Therapy, 41*, 1251–1269.
- Marks, I. M., & Matthews, A. M. (1979). Brief standard self-rating for phobic patients. *Behaviour Research and Therapy, 17*, 263–267.
- Merrill, K. A., Tolbert, V. E., & Wade, W. A. (2003). Effectiveness of cognitive therapy for depression in a community mental health center: A benchmarking study. *Journal of Consulting and Clinical Psychology, 71*, 404–409.
- Morgenstern, J., Blanchard, K. A., Morgan, T. J., Labouvie, E., & Hayaki, J. (2001). Testing the effectiveness of cognitive-behavioral treatment for substance abuse in a community setting: Within treatment and posttreatment findings. *Journal of Consulting and Clinical Psychology, 69*, 1007–1017.
- Mueller, R., Lambert, M. J., & Burlingame, G. (1998). The Outcome Questionnaire: A confirmatory factor analysis. *Journal of Personality Assessment, 70*, 248–262.
- Otto, M. W., Pollack, M. H., & Sabatino, S. A. (1996). Maintenance of remission following cognitive behavior therapy for panic disorder: Possible deleterious effects of concurrent medication treatment. *Behavior Therapy, 27*, 473–482.
- Otto, M. W., Pollack, M. H., Sachs, G. S., Teiter, S. R., Meltzer-Brody, S., & Rosenbaum, J. F. (1993). Discontinuation of benzodiazepine treatment: Efficacy of cognitive-behavioral therapy for patients with panic disorder. *American Journal of Psychiatry, 150*, 1485–1490.
- Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy and Research, 23*, 535–548.
- Persons, J. B., & Silberschatz, G. (1998). Are the results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66*, 126–135.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Application and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Seggar, L. B., Lambert, M. J., & Hansen, N. B. (2002). Assessing clinical significance: Application to the Beck Depression Inventory. *Behavior Therapy, 33*, 253–269.
- Shear, M. K., & Maser, J. D. (1994). Standardized assessment for panic disorder research: A conference report. *Archives of General Psychiatry, 51*, 346–354.
- Shear, M. K., Ricci, P., Williams, J., Frank, E., Grochocinski, V., Vander Bilt, J., et al. (2001). Reliability and validity of the Panic Disorder Severity Scale: Replication and extension. *Journal of Psychiatric Research, 35*, 293–296.
- Sheidow, A. J., Bradford, W. D., Henggeler, S. W., Rowland, M. D., Halliday-Boykins, C., Schoenwald, S. K., et al. (2004). Treatment costs for youths receiving multisystemic therapy or hospitalization after a psychiatric crisis. *Psychiatric Services, 55*, 548–554.
- Silverman, W. H. (1996). Cookbooks, manuals, and paint-by-numbers psychotherapy in the 90s. *Psychotherapy, 33*, 207–215.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 24*, 323–355.
- Spiegel, D. A., Bruce, T. J., Gregg, S. F., & Nuzzarelo, A. (1994). Does cognitive behavior therapy assist slow-taper alprazolam discontinuation in panic disorder? *American Journal of Psychiatry, 151*, 876–881.
- Street, L. L., Niederhe, G., & Lebowitz, B. D. (2000). Toward greater public health relevance for psychotherapeutic intervention research: An NIMH workshop report. *Clinical Psychology: Science and Practice, 7*, 127–137.
- Tuschen-Caffier, B., Pook, M., & Frank, M. (2001). Evaluation of manual-based cognitive-behavioral therapy for bulimia nervosa in a service setting. *Behaviour Research and Therapy, 39*, 299–308.
- Umphress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15*, 40–55.
- Wade, W. A., Treat, T. A., & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. *Journal of Consulting and Clinical Psychology, 66*, 231–239.
- Warren, R., & Thomas, J. C. (2001). Cognitive-behavior therapy of obsessive-compulsive disorder in private practice: An effectiveness study. *Journal of Anxiety Disorders, 15*, 277–285.
- Weisz, J. R., & Addis, M. E. (in press). The research-practice tango and other choreographic challenges: Using and testing evidence-based psychotherapies in clinical care settings. In C. D. Goodheart, A. E. Kazdin, & R. J. Sternberg (Eds.), *Evidence-based psychotherapy: Where prac-*

- tice and research meet.* Washington, DC: American Psychological Association.
- Weisz, J. R., Southam-Gerow, M. A., Gordis, E. B., & Connor-Smith, J. (2003). Primary and secondary control enhancement training for youth depression: Applying the deployment-focused model of treatment development and testing. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 165–183). New York: Guilford Press.
- Wells, K., Sherbourne, C., Schoenbaum, M., Ettner, S., Duan, N., Miranda, J., et al. (2004). Five-year impact of quality improvement for depression: Results of a group-level randomized controlled trial. *Archives of General Psychiatry*, *61*, 378–386.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *69*, 875–899.
- Westra, H. A., Stewart, S. H., & Conrad, B. E. (2002). Naturalistic manner of benzodiazepine use and cognitive behavioral therapy outcome in panic disorder with agoraphobia. *Journal of Anxiety Disorders*, *16*, 233–246.
- Willet, J. B., Singer, J. D., & Martin, N. C. (1998). The design and analysis of longitudinal studies of development and psychopathology in context: Statistical methods and methodological recommendations. *Development and Psychopathology*, *10*, 395–426.
- Wilson, G. T. (1996). Manual-based treatments: The clinical application of research findings. *Behaviour Research and Therapy*, *34*, 295–314.

Received October 1, 2004

Revision received September 15, 2005

Accepted September 23, 2005 ■