Therapists, Therapist Variables, and Cognitive–Behavioral Therapy Outcome in a Multicenter Trial for Panic Disorder

Jonathan D. Huppert
Boston University and New York-Presbyterian Hospital–
Weill College of Medicine at Cornell University

Lynn F. Bufka and David H. Barlow
Boston University

Jack M. Gorman
Columbia University and Long Island
Jewish/Hillside Medical Center

M. Katherine Shear
University of Pittsburgh School of Medicine

Scott W. Woods
Yale University School of Medicine

The relationship between therapists and treatment outcome was examined in 14 highly trained therapists who participated in the Multicenter Collaborative Study for the Treatment of Panic Disorder. Overall, therapists yielded positive outcomes in their caseloads; yet, therapists significantly differed in the magnitude of change among caseloads. Effect sizes for therapist impact on outcome measures varied from 0% to 18%. Overall experience in conducting psychotherapy was related to outcome on some measures, whereas age, gender, gender match, and experience with cognitive–behavioral therapy (CBT) were not. Therapists with above- and below-average outcomes were rated similarly on measures of adherence and competency. The results suggest that therapists make a contribution to outcome in CBT for panic disorder, even when patients are relatively uniform, treatment is structured, and outcome is positive. Implications for future clinical outcome studies and for training clinicians are discussed.

The therapist's contribution to outcome in therapy has been a focus of considerable interest (see the special section of *Clinical Psychology: Science and Practice*, Kazdin, 1997; Beutler, Machado, & Neufeldt, 1994), particularly since the establishment of a number of empirically supported psychological interventions for a variety of disorders (Barlow, 1994, 1996; Nathan & Gorman, 1998; see also the special section of the *Journal of Consulting and Clinical Psychology*, Kendall & Chambless, 1998). A number of issues have been addressed under the rubric of therapist variables or therapist factors that influence outcome in psychotherapy. One

Jonathan D. Huppert, Center for Anxiety and Related Disorders, Boston University, and Department of Psychology in Psychiatry, New York-Presbyterian Hospital–Weill College of Medicine at Cornell University; Lynn F. Bufka and David H. Barlow, Center for Anxiety and Related Disorders, Boston University; Jack M. Gorman, Department of Psychiatry, Columbia University, and Department of Psychiatry, Long Island Jewish/Hillside Medical Center; M. Katherine Shear, Department of Psychiatry, University of Pittsburgh School of Medicine; Scott W. Woods, Department of Psychiatry, Yale University School of Medicine.

Correspondence concerning this article should be addressed to Jonathan D. Huppert, who is now at the Center for Treatment and Study of Anxiety, University of Pennsylvania, 3535 Market Street, 6th Floor, Philadelphia, Pennsylvania 19104. Electronic mail may be sent to huppert@mail.med.upenn.edu.

question is whether certain therapists have better outcomes overall or whether particular therapists have better outcomes with specific patient populations or disorders (i.e., Luborsky et al., 1986; Luborsky, McLellan, Diguer, Woody, & Seligman, 1997). In addition, researchers have conducted analyses evaluating what specific factors or variables contribute to differences in therapists' outcomes (e.g., Lafferty, Beutler, & Crago, 1989). Such factors include demographic characteristics (i.e., age, gender, race, religion), training characteristics (i.e., degree, training, years of experience), beliefs (orientation), and personality characteristics.

Most studies examining the relationship between the therapist's age and treatment outcome have been based on uncontrolled, naturalistic studies. These studies have tended to report that the therapist's age is not significantly related to therapy outcome (D. F. Beck, 1988; Greenspan & Kulish, 1985). Other studies have focused on how therapists' and clients' ages interact to affect treatment outcomes (see Atkinson & Schein, 1986, for an extensive review of this literature). Although some studies have reported a small relationship between therapist and client age similarity and treatment outcome (e.g., Dembo, Ikle, & Ciarlo, 1983), the general conclusion remains that the difference between therapist and client age is a weak predictor of therapy outcome (Beutler et al., 1994). This conclusion has been supported by clinical studies with samples of older adults (Thompson, Gallagher, & Breckenridge, 1987) and samples of children and adolescents (Weisz, Weiss, Han, Granger, & Morton, 1995), where therapists' and clients' ages are expected to differ significantly.

Studies examining the effects of the therapist's gender on treatment outcomes have produced rather inconsistent results (Beutler

et al., 1994). Bowman (1993) reviewed the literature and concluded that the data currently support three competing hypotheses about the effects of therapist gender on outcomes in therapy: (a) female therapists are more effective than male therapists (e.g., Fisher, 1989), (b) better outcomes are produced when therapists and clients are of the same gender (Luborsky, Auerbach, Chandler, Cohen, & Bachrach, 1971; Persons, Persons, & Newmark, 1974), and (c) therapist gender is not significantly related to therapy outcomes (e.g., Sexton & Whiston, 1991). In addition, therapist gender or therapist–patient gender match did not contribute to outcome in the National Institute of Mental Health (NIMH) Treatment of Depression Collaborative Research Project (TDCRP; Zlotnick, Elkin, & Shea, 1998). Overall, therapists' skill level and experience are usually considered more robust predictors of outcomes than is therapists' gender (Beutler et al., 1994).

Therapists' skill level, techniques, and experience have been investigated in relationship to treatment outcome. This literature examining individual therapists' contributions to treatment outcome has generally concluded that better outcomes are seen with some therapists compared with others (Luborsky, McLellan, Woody, O'Brien, & Auerbach, 1985; Luborsky et al., 1986). In fact, it has been argued that individual therapists' contributions to outcome have been the dominant factor in demonstrating differences between different treatments (Crits-Christoph & Mintz, 1991; Martindale, 1978; Ricks, 1974). Also, some studies have reported that more experience does seem to lead to better outcomes, although the differences have been modest at best (Crits-Christoph et al., 1991; Stein & Lambert, 1995).

The original studies examining therapist effects did not administer standardized treatments (e.g., Ricks, 1974). However, a significant, positive effect on treatment outcome has been found to occur with therapists' greater adherence to treatment manuals (Luborsky et al., 1985, 1997). In a meta-analysis of 27 separate treatment groups, use of a treatment manual was associated with smaller differences between therapists, whereas not using a treatment manual was associated with larger between-therapist differences (Crits-Christoph et al., 1991). Additionally, the authors found that more experienced therapists had smaller between-therapist differences than did less experienced therapists, though wide variability existed across studies. In another recent examination of therapist effects, Luborsky et al. (1997) examined the performance of 22 therapists across seven manualized treatment outcome studies. They concluded that therapists varied greatly in the mean amount of improvement observed in their clients (from slightly negative to 80% positive) but that therapists who excelled in one sample excelled in others as well. Finally, they concluded that more information is needed to understand the variables that influence differences in patient outcome. Large clinical trials may provide the appropriate data to examine these issues further.

Although examining therapist variables is not the primary purpose of large clinical trials, such studies have the advantage of involving a reasonably large number of therapists, administering treatments in a controlled fashion after similar training (usually with treatment manuals), and studying similar patient populations (i.e., depression, panic). For example, Blatt, Sanislow, Zuroff, and Pilkonis (1996) examined therapist variables in the NIMH's TDCRP (Elkin, Parloff, Hadley, & Autry, 1985). Therapists were categorized into more effective, moderately effective, and least effective groups, and significant differences in therapists were found in the TDCRP, regardless of treatment and research site. Blatt et al. concluded that neither the demographics of therapists nor the attitudes of therapists concerning the treatment and etiology of depression (biological vs. psychological) appeared to influence effectiveness. However, therapists' reported prestudy use of psychosocial, biological, or combined treatments was related to outcome. Clinicians who reported using more frequent psychosocial interventions alone were more effective in the TDCRP. Generalizing from her experience with outcome and therapist effects in the TDCRP, Elkin (1999) noted that evaluating therapist variability is an important, yet often underemphasized, aspect in determining treatment outcome and she espoused that therapist characteristics must be evaluated in treatment outcome studies.

In this article, we explore the role of therapist age, gender, treatment orientation, experience, and specific experience with cognitive–behavioral therapy (CBT) on treatment outcome. We also evaluate the global impact of therapists on treatment outcome in a randomized clinical trial of the treatment of panic disorder by using an empirically supported CBT treatment and multiple outcome measures. The data from the Multicenter Collaborative Study for the Treatment of Panic Disorder (MCSTPD; Barlow, Gorman, Shear, & Woods, 2000) were used to examine therapist effects on treatment outcome when conducting the current psychotherapy of choice (i.e., CBT).

Method

Treatment Study

Patients were randomly assigned to one of five groups: panic control treatment (CBT only; hereinafter CBT-O), panic control treatment plus placebo (CBT-PL), panic control treatment plus imipramine (CBT-IMI), imipramine alone, and placebo alone. For purposes of the current analyses, all CBT groups (e.g., CBT-O, CBT-PL, and CBT-IMI) were combined and analyzed together. (For more details, see Barlow et al., 2000.)

Therapists

From among doctoral-level therapists participating in the MCSTPD who had data from four or more patients, 14 (7 men and 7 women, 13 of whom were psychologists and 1 of whom was a psychiatrist) were identified for analysis. Therapists were doctoral-level employees of the anxiety clinics at each site. The average age of the therapists was 35.7 years ($SD = 6.3$; range $= 27–45$). Therapists varied in experience conducting general psychotherapy from 2 to 20 years ($M = 8.9$, $SD = 5.6$) and in conducting CBT from 1 to 18 years ($M = 5.9$, $SD = 5.1$ years). Nine therapists described their orientation as primarily CBT, whereas 5 described themselves as other (i.e., eclectic or psychodynamic). All therapists were trained to competency and certified in conducting panic control treatment (Craske, Barlow, & Meadows, 2000) prior to participating in the active phase of treatment. Training consisted of participating in a didactic seminar, viewing videotaped sessions of a case, and treating patients under close supervision. Extensive feedback and supervision were given during training, and therapists needed to reach a high level of rated adherence and competence prior to being certified. Supervision continued during biweekly conference calls throughout the study. After the clinical trial began, no therapist demonstrated further improvement across patient outcomes over time (Huppert et al., 1997), suggesting that therapists were trained sufficiently from the outset.

Patients

Of the total sample of 312 individuals treated for panic disorder, 205 patients were randomized to CBT. Of those, 22 either dropped out prior to

the first session ($n = 10$), were treated by a therapist who treated fewer than 4 patients in the study ($n = 4$), or were missing substantial data ($n = 8$). Therefore, we examined data from 183 patients (70 men, 113 women) who were assigned to one of the CBT arms of the trial (67 CBT-O, 59 CBT-IMI, 57 CBT-PL). Patients were distributed across four sites that are specialty centers for the treatment of anxiety: two known for psychosocial treatments and two known for psychopharmacological treatments. Patients were predominantly Caucasian (166 Caucasian, 17 African American, Hispanic, or Other), and their ages ranged from 19 to 65 years ($M = 36$, $SD = 11.0$). Patients had been diagnosed with panic disorder without agoraphobia (PDA) or with low levels of agoraphobia by trained diagnosticians using the Anxiety Disorders Interview Schedule—Revised (ADIS-R; DiNardo & Barlow, 1988). All patients had at least one full- or limited-symptom panic attack in the 2 weeks preceding the evaluation. Patients were excluded from the study if they had ever had a psychotic disorder or if they were currently suffering from significant medical illnesses, suicidality, or significant substance abuse. Patients were excluded for contraindications to either CBT or medication treatment, prior nonresponse to CBT or imipramine, or concurrent treatment or disability claims. Patients were evaluated with a number of independent observer ratings and self-report measures, to be discussed below.

Therapist Characteristic Measures

Data regarding the age, gender, self-identified theoretical orientation, years of experience doing therapy, and years of experience with CBT techniques were collected from therapists through circulation of a questionnaire at each treatment site (see Table 1). The principal investigator from each site completed the questionnaire with the assistance of therapists at each site. The investigators were not aware of individual therapists' caseload outcomes at the time of completion of the questionnaire. Therapists who were identified as having a CBT orientation were younger than other therapists, $t(12) = 2.61$, $p < .05$, but had no less experience in therapy overall, $t(12) = 1.25$, $p > .10$, and had somewhat more experience with CBT, $t(12) = 1.96$, $p < .10$.

Table 1
Descriptives of Individual Therapists

Site	Gender	Age (years)	Years of CBT	Years of therapy	Orientation
W	Female	42	3	10	Other[a]
W	Male	29	8	8	CBT
X	Female	27	4	4	CBT
X	Female	45	4	15	Other
X	Male	40	16	16	Other
X	Female	35	.2	4	Other
X	Male	43	18	20	CBT
Y	Female	29	6	6	CBT
Y	Male	32	5	5	CBT
Y	Male	29	2	2	CBT
Y	Male	30	6	6	CBT
Z	Female	41	3	3	CBT
Z	Female	41	1	14	Other
Z	Male	37	5	11	CBT
Total	7 Male				9 CBT
	7 Female				5 Other
Average		35.7	5.9	8.9	

Note. In order to keep the identities of the therapists confidential, the order of therapists here do not coincide with therapist numbers in other tables. CBT = cognitive–behavioral therapy.
[a] Other includes psychodynamic or eclectic orientation.

As part of the protocol for the study, adherence and competency measures were collected for a random sample of patients and for a random sample of sessions per patient. Adherence was rated by how thoroughly specific concepts and techniques were addressed in each session on the basis of a 7-point scale from 1 (*not done*) to 7 (*extensively covered*), with 3 being rated as *some discussion* and 5 as *considerable discussion*. Competency was globally rated for each session, with a single rating on a 7-point scale from 1 (*clearly inadequate*) to 7 (*excellent*), with 3 rated as *fair* and 5 as *good*. A therapist was rated as excellent if he or she had a warm, supportive, collaborative, Socratic style and was able to articulate the concepts clearly, making them personally relevant to the patient. These measures were developed for use with this trial and do not have external data on validity. However, all adherence and competency raters were carefully trained to a high level of reliability prior to rating sessions. A random sample of sessions and patients evaluated by raters unaware of the therapist's identity and treatment condition revealed high levels of adherence and competency during treatment throughout the study (adherence mean = 5.72, $SD = 0.70$; competency mean = 5.59, $SD = 1.06$; scales range from 1 to 7).

Patient Outcome Measures

A number of panic-related measures were collected at pretreatment and post–acute-treatment phase (i.e., after 12 weeks of treatment). Trained clinicians, acting as independent evaluators and unaware of patient treatment assignment, evaluated patients by using the following measures: ADIS-R, Panic Disorder Severity Scale (PDSS; Shear et al., 1997), Clinical Global Impressions Scale anchored for panic disorder (CGI; Guy, 1976), Hamilton Anxiety Scale (HAS; Hamilton, 1959), Hamilton Depression Scale (HDS; Hamilton, 1960), and the severity of anticipatory anxiety (rated as a separate item collected during the ADIS-R interview; AntAnx). Self-report measures included the Anxiety Sensitivity Index (ASI; Peterson & Reiss, 1992) and the Subjective Symptom Scale (SSS; Hafner & Marks, 1976). Frequency of panic attacks (FPA) was also collected through self-report (from the PDSS). A categorical variable, response status, was determined by masked raters by a patient achieving a score of at least 2 (*much improved*) and being rated as 3 (*mild*) or less on 7-point scales from the CGI. Patients who discontinued treatment prior to completing five sessions were categorized separately as dropouts. Participants who had missing data were still included in analyses in which their data were available. However, because follow-up data were not collected on dropouts, all posttreatment measures included only patients who completed treatment.

Data Analyses

Effect sizes were calculated for the effect of the therapist on each individual outcome measure (Crits-Christoph et al., 1991). For each outcome measure, analyses of covariance (ANCOVAs) were conducted, with the posttreatment score as the dependent variable, therapist as the independent variable, and pretreatment score as a covariate. From each ANCOVA, an estimate of the variance for therapist was calculated. This variance was divided by the total variance in the dependent variable, thus calculating the percent variance accounted for by therapist, that is, the therapist effect size (Crits-Christoph et al., 1991). In addition, because we found a main effect for treatment assignment in previous analyses (Barlow et al., 2000), we again conducted ANCOVAs using the posttreatment measures scores as the dependent variables, therapists as independent variables, and pretreatment scores and treatment assignment as covariates.

Residualized change scores were then calculated for each measure by regressing the pretreatment score of each measure onto posttreatment scores. The standardized residual was then calculated and used as the change score for each patient/measure. Next, the residualized change scores were analyzed with principal-components analysis to determine

whether all measures had a common underlying factor. All measures loaded at least .60 on the only factor extracted (eigenvalue > 1), with the PDSS loading .90 on this factor. Because the PDSS, a measure developed to comprehensively evaluate panic disorder symptoms (Shear et al., 1997), was so highly loaded on this factor, we decided to use this measure as our major indicator of change. The PDSS was significantly correlated with all measures, ranging from .77 (with the CGI) to .47 (with the ASI). Overall, all measures correlated at least .40 with each of the other measures.

A number of analyses were conducted to evaluate the relationship between therapist characteristics and outcome measures. Pearson correlations were calculated between the residualized change scores and therapist characteristics. Kendall's tau was used for categorical variables (gender, theoretical orientation). We also conducted analyses of variance (ANOVAs) to examine patient–therapist match for gender. In addition, therapists were compared on adherence and competency measures to determine whether individual therapists differed in rated adherence and competency.

The 14 therapists were subdivided into three groups on the basis of treatment outcome: Six were allocated to the above average group (A therapists), 4 were determined to be in the average group (B therapists), and 4 were placed in the below average group (C therapists). Assignments were made by determining the number of measures in which a therapist was ranked in the top half, according to residualized change scores. There were 6 therapists who fell clearly in the A therapists group (measured in the top 66% on 80% or more of the measures) and four who clearly fell in the C therapists group (measured below average on 100% of the measures). The remaining four therapists were placed in the B therapists group. An omnibus chi-square analysis determined the likelihood that groups of therapists differed on response status for both the full sample ($N = 183$; i.e., intent to treat; counting dropouts as nonresponders) and the patients who completed treatment ($n = 154$). Follow-up chi-square tests were then conducted to determine whether A therapists differed from C therapists and whether B therapists differed from A therapists or C therapists. Therapist differences were also calculated for pretreatment measures as well as for adherence, competency, and therapist characteristics.

In addition, the numbers of patients who improved, did not change, or deteriorated were calculated according to each outcome measure. Differences between A and C therapists were then examined through Fisher's exact test. To test whether A and C therapists differed in average magnitude of change in these patients who had improved, we conducted ANCOVAs using postacute measures as the dependent variables, A or C therapist assignment as the independent variable, and pretreatment scores as the covariate.

Results

Therapist effect sizes varied greatly depending on which measure of outcome was used. The PDSS showed a modest effect size,

8%, $F(13, 138) = 1.87$, $p = .04$. The ASI showed the greatest therapist effect, 18%, $F(13, 120) = 3.03$, $p = .001$; followed by the ADIS–R, 15%, $F(13, 133) = 2.79$, $p = .002$; and HAS, 8%, $F(13, 137) = 1.85$, $p = .04$. The other measures showed minimal therapist effects: HDS, 4%, $F(13, 137) = 1.48$, $p = .13$; SSS, 4%, $F(13, 117) = 1.36$, $p = .19$; CGI, 3%, $F(13, 138) = 1.36$, $p = .19$; AntAnx, 2%, $F(13, 134) = 1.23$, $p = .26$; and FPA, 1%, $F(13, 135) = 1.10$, $p = .36$. F values did not change more than 0.02 when covarying out treatment assignment in addition to pretreatment scores in our ANCOVAs, meaning that treatment assignment was not influencing our results. Therefore, we did not include treatment assignment in further analyses.

Table 2 presents the results of correlations between therapists' characteristics and average residualized change scores per caseload. PDSS change scores were not related to any therapist characteristics. In contrast, significant relationships were found between average change in caseloads based on some other outcome measures and years of experience conducting therapy (ASI, $r = -.72$, $p < .01$; ADIS–R, $r = -.57$, $p < .05$), and age (ADIS–R, $r = -.64$, $p < .05$). Correlations were negative because residualized change scores were used as the outcome variables. Visual examination of scatterplots confirmed that these correlations were not due to outliers. Thus, therapists with more experience in therapy in general were more likely to have patients' anxiety sensitivity decrease, and older therapists and therapists with more experience were associated with more change in overall panic disorder severity, as rated on the ADIS–R. An examination of the therapist variables showed that there was not a significant relationship between years of CBT experience and years of therapy experience. No differences in outcome were found when examining the effects of therapist–patient match or mismatch for gender (i.e., female–female, male–male, male–female, female–male; ANOVAs, all $Fs < 1$, $ps > .3$).

Therapists also differed in adherence and competency, although all exceeded the standard acceptable for the study. A total of 330 sessions were rated on adherence, and 526 were rated on competency. ANOVAs showed significant differences between individual therapists on adherence, $F(12, 332) = 9.07$, $p < .001$, and on competency, $F(13, 515) = 17.19$, $p < .001$. Post hoc analysis revealed multiple differences among therapists.

Table 3 visually demonstrates the number of times each therapist was given each ranking. Rows (i.e., therapists) are ordered to

Table 2

Correlations Between Therapist Variables and Average Change in Therapist Caseloads on Each Outcome Measure (N = 14)

Therapist variable	PDSS	CGI	ADIS–R	FPA	HDS	HAS	SSS	AntAnx	ASI
Age (years)	−.221	−.340	−.643*	−.251	−.092	−.373	−.332	−.205	−.378
Years of experience	−.399	−.348	−.566*	−.318	−.111	−.503	−.507	−.212	−.723**
Years of CBT experience	−.217	−.087	.062	.009	.061	.096	−.099	.032	−.515
Gender	.165	.314	.135	.135	.105	.285	.195	.344	.015
Orientation	−.172	−.203	−.453	−.328	−.141	−.453	−.047	−.203	.016

Note. Pearson correlations were calculated for age, years of experience, years of cognitive–behavioral therapy (CBT) experience. Kendall's tau was calculated for orientation and gender. PDSS = Panic Disorder Severity Scale; CGI = Clinical Global Impressions Scale; ADIS–R = Anxiety Disorders Interview Schedule—Revised; FPA = frequency of panic attacks; HDS = Hamilton Depression Scale; HAS = Hamilton Anxiety Scale; SSS = Subjective Symptom Scale; AntAnx = anticipatory anxiety rated during the ADIS–R interview; ASI = Anxiety Sensitivity Index.
* $p < .05$. ** $p < .01$.

Table 3
Number of Times Each Therapist Received Each Ranking

Therapist	Ranking													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Therapist grouping A														
1	4	3	2											
3	2		2	1	1	2		1						
7	1	2	2	1	1		2							
6	1				2	3		2	1					
4				3	3	1				2				
10			1	1	1		1	2	1	2				
Therapist grouping B														
2		2		1		1	2	1			1	1		
5	1	2	2		1		1						1	1
9			1				1	1	2		2	2		
12			1			2	2			1	1	1		
Therapist grouping C														
11								2	2	2	3			
8							1	1	2	3	1	1		
14							1	1			1	5	1	
13												2	7	

Note. The top 5 therapists accounted for the large majority of above average ranks on individual measures, whereas the bottom 4 therapists never ranked in the top 6 on any individual measure. The underscored number represents the ranking on the Panic Disorder Severity Scale.

demonstrate the clear difference in efficacy among different therapists.

Results of post hoc analyses examining which therapists differed from each other on the residualized measures showed that Therapists 1 and 3 were more effective with their caseloads than other therapists according to a number of outcome measures, whereas Therapists 13 and 14 had less than average change on a number of measures.

It is important to note that, on average, all therapists but one obtained marked improvement in their caseloads. The average percent change (pretest minus posttreatment) divided by pretest, across all measures was 49% (SD = 37%). The PDSS, one of the most comprehensive measures of change in severity of PDA, showed an average change of 58% (SD = 31%). Most patients had lower posttreatment scores for most measures, compared with pretest scores. Patients of only one of the therapists, who saw just four patients during the study, showed a deterioration in average posttest scores compared with pretest scores. Thus, for most therapists, the discussion concerns whose patients improved most on average. For example, in examining change by groups, we found that the A therapists had 67% improvement on the PDSS, whereas C therapists had 48% improvement. The A therapists had an average decrease of 3 points (59% improvement) on the ADIS–R, whereas the C therapists had an average decrease of 2 points (42% improvement). Finally, on the ASI, A therapists had an average decrease of 20 points (58% improvement), whereas C therapists had an average decrease of 13 (39% improvement). In examining individual therapists, we find that even the least effective therapist's caseload had positive change, on average, on 5 out of 9 measures, no change on 2, including the most important outcome

measure (PDSS), and negative change on 2 (HDS and SSS). However, that therapist treated only four patients (one of whom improved on 8 of 9 measures and stayed the same on the other measure; SSS).

Using Jacobson's reliable change index, with two standard deviations below the pretreatment mean for the criteria for clinical significance, we were able to calculate the number of patients who met full criteria for reliable and clinically significant change (Jacobson & Truax, 1991) on the PDSS. For A therapists, 66% of their caseload met the criteria (n = 33 responders), and for C therapists, 45% of their cases met the criteria (n = 27 responders). Because extensive normative data exist for the ASI (Peterson & Reiss, 1992; n = 4,517, M = 19.01, SD = 9.11) and normative data have not been published for the PDSS and ADIS–R to date, we decided to examine the reliability of change (Jacobson & Truax, 1991) and the clinical significance of change (Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999) on the ASI. Using Kendall's method of clinical equivalency, we found that C therapists' patients' posttreatment scores were clinically equivalent to those of controls, to the normal range, $t_{CE}(4568)$ = 6.93, p < .001, and did not differ from nondiagnosed individuals in posttreatment scores, $t_{trad}(4568)$ = 0.29, p > .35, thereby showing that the patients were clinically comparable to the normal range on the ASI. We found that patients treated by the A therapists had ASI scores below the normal distribution, approaching one standard deviation below the norm, $t_{CE}(4561)$ = 1.44, p < .10, and their scores were significantly lower than those of the undiagnosed control sample, $t_{rad}(4561)$ = 5.30, p < .001. This suggests that patients treated by A therapists had substantially lower scores than the nondiagnosed population. Using Jacobson's criteria for reliable

change, we found that 52% of the patients seen by A therapists had reliable and significant change on the ASI, and 35% of patients treated by C therapists had reliable and significant change.

Omnibus chi-square tests revealed significant differences among A, B, and C therapists in response status (see Table 4), $\chi^2(2, N = 183) = 8.75, p = .013$. Note that response status is a variable that was not considered for the rankings and therefore could have yielded different results than the other measures (discussed later). For the completer analysis, A therapists had 5 nonresponders and 46 responders, B therapists had 7 nonresponders and 36 responders, and C therapists had 19 nonresponders and 41 responders. There was no difference between groups in the number of dropouts (A therapists = 12, B therapists = 7, C therapists = 10).

Follow-up analyses revealed that the differences in response status were attributable to A and C therapists. A therapists had a higher percentage of patients considered responders than did C therapists, $\chi^2(1, N = 133) = 7.78, p = .005$. Differences between A and B therapists and between B and C therapists did not approach significance, A–B: $\chi^2(1, N = 113) = 1.39, p = .499$; B–C: $\chi^2(1, N = 120) = 3.15, p = .207$.

Differences between A and C therapists were explored by looking at pretreatment scores on patient measures to determine whether one group had more severely ill patients or whether most improvement could be accounted for by any specific measure. To correct for the number of analyses conducted, we used .01 for the alpha level. There were no differences between A and C therapists' caseloads on pretreatment PDSS, $t(129) = 0.73, p = .467$, or any other measures, HAS: $t(129) = 2.18, p = .031$; HDS: $t(129) = 1.98, p = .05$; all other ts < 1.6, ps > .10.

Table 4
Response Status and Number of Dropouts for Each Therapist

Therapist grouping and therapist no.	Response status			
	Responders	Nonresponders	Dropouts	Total
Therapist grouping A				
1	3		1	4
3	6		1	7
7	5		1	6
6	11	3	7	21
4	11	2		13
10	10		2	12
Total	46	5	12	63
Therapist grouping B				
2	9		5	14
5	4			4
9	13	6	2	21
12	10	1		11
Total	36	7	7	50
Therapist grouping C				
11	19	11	3	33
8	13	3	3	19
14	7	3	4	14
13	2	2		4
Total	41	19	10	70
Total	123	31	29	183

A and C therapists differed in the number of patients that showed no improvement or worsened according to a number of measures, though not on the PDSS (see Table 5; ADIS–R Fisher's exact test, $p = .01$; AntAnx Fisher's exact test, $p = .049$; ASI Fisher's exact test, $p = .019$; CGI Fisher's exact test, $p = .009$). Although numbers were too small to detect differences in the number of patients that deteriorated according to the various measures, some interesting contrasts are presented in Table 5. The A therapists appeared to have fewer patients worsen or show no change. To test whether differences between A and C therapists were solely due to patients who had deteriorated or did not change, we examined differences between A and C therapists on each measure by selecting only those patients who improved on the measure. Results revealed that A and C therapists continued to differ on a number of outcome measures including PDSS, $F(1, 101) = 8.96, p = .003$; AntAnx, $F(1, 91) = 7.00, p = .010$; ASI, $F(1, 80) = 7.49, p = .008$; FPA, $F(1, 76) = 11.22, p = .001$; HAS, $F(1, 90) = 6.29, p = .014$; HDS, $F(1, 80) = 8.75, p = .004$; and SSS, $F(1, 79) = 6.48, p = .013$. Differences between the two groups for ADIS–R, $F(1, 92) = 3.79, p = .056$, and CGI, $F(1, 90) = 3.08, p = .083$, approached significance.

No differences were found between A and C therapists in orientation (Group A: 4 CBT, 2 other, Group C: 3 CBT, 1 other; Fisher's exact test $p = 1.0$), therapist gender (Group A: 4 women, 2 men, Group C: 2 women, 2 men; Fisher's exact test $p = 1.0$), age (A therapists $M = 37.5$ years, $SD = 6.9$, C therapists $M = 35.5$ years, $SD = 7.0$; $t[8] = 0.447, p = .667$), years of therapy experience (A therapists $M = 11.5, SD = 6.2$, C therapists $M = 5.2, SD = 3.6$; $t[8] = 1.81, p = .107$), and years of CBT experience (A therapists $M = 6.2, SD = 3.0$, C therapists $M = 3.5, SD = 1.7$; $t[8] = .85, p = .419$).

No differences in adherence were found between the three groups, $F(2, 342) = .62, p > .5$, but groups did differ in competence, $F(2, 526) = 9.04, p < .001$. Follow-up tests revealed that competency ratings of B therapists were significantly higher than those of A and C therapists (ps < .01), but A and C therapists did not differ in competency ratings from each other.

Results also confirmed that treatment assignments did not differ between A and C therapists, $\chi^2(2, N = 133) = 3.01, p = .222$. A therapists had 23 patients in CBT-IMI, 20 in CBT-PLA, and 20 in CBT, whereas C therapists had 16 patients in CBT-IMI, 26 patients in CBT-PLA, and 28 patients in CBT alone.

Site differences between therapists were not significant (ps > .4; A therapists: 3 Site W, 2 Site Y, 1 Site Z; B therapists: 2 Site W, 1 Site X, 1 Site Y; C therapists: 1 Site Y, 1 Site X, 2 Site Z).

Discussion

We assessed the relationship between several therapist characteristics and treatment outcome by examining the relationship between commonly observed characteristics (i.e., experience, age, sex, and orientation) of 14 therapists who were involved in the MCSTPD and outcome measures. Although there were no effects for experience on our primary outcome measure, patients who were seen by more experienced therapists showed greater improvement than did patients seen by less experienced therapists on two measures that are important for CBT conceptualizations of panic, including anxiety sensitivity (see Taylor, 1999). Experience was related to outcome more when it was defined as overall years

Table 5

Summary of Change Differences Between A and C Therapists

Measure and group	Negative change	Negative or no change	Positive change
PDSS			
A	1	1	62
C	4	5	65
CGI			
A*	1	3	60
C	1	14	56
ADIS–R			
A*	0	1	62
C	1	10	60
FPA			
A	0	13	50
C	5	16	54
HDS			
A	7	9	54
C	12	18	52
HAS			
A	5	7	56
C	9	10	60
SSS			
A	1	4	59
C	8	10	60
AntAnx			
A*	1	3	60
C	3	11	59
ASI			
A*	2	2	61
C	9	11	59

Note. PDSS = Panic Disorder Severity Scale; CGI = Clinical Global Impressions Scale anchored for panic disorder; ADIS–R = Anxiety Disorders Interview Schedule—Revised; FPA = frequency of panic attacks; HDS = Hamilton Depression Scale; HAS = Hamilton Anxiety Scale; SSS = Subjective Symptom Scale; AntAnx = anticipatory anxiety rated during the ADIS–R interview; ASI = Anxiety Sensitivity Index.
* $p < .05$ (see text for statistical results).

of practicing psychotherapy than when it was defined as years of practicing CBT. This finding supports the often elusive notion that experience may be related to outcome. In a randomized study that examined a manualized, stage-driven CBT treatment there are many reasons to believe that experience would not matter. Therapists had all received rigorous training in the protocol, which although still flexible and individualized for patients, was perhaps somewhat more structured than other CBT treatments, such as CBT for depression (A. T. Beck, Rush, Shaw, & Emery, 1985). Supervision was continuously provided and adherence and competency ratings were high for all therapists. However, despite these factors, experience was associated with improved outcome, though only on some measures. It is possible that a deeper understanding of general CBT principles or more skill in developing a therapeutic alliance (Strupp & Anderson, 1997) is gained through more experience conducting psychotherapy. Further exploration of Beutler's (1997) contention that a more precise definition of experience (e.g., total number of patient hours) might help further differentiate the experience effect.

The effects of other therapist characteristics, such as theoretical orientation, therapist gender, and age were not significantly related to outcome. This finding is similar to the findings of most studies of the impact of these therapist characteristics on outcome (e.g.,

Beutler et al., 1994). In addition, above- and below-average therapists were not differentiated by adherence or competency measures, implying that technique itself is not the only important factor related to outcome (cf. Shaw et al., 1999). We suggest a number of reasons why adherence and competency may not have been related to outcome. First, it is possible that while remaining within the boundaries of the protocol, a more flexible approach, as opposed to more rigid adherence, could be related to better outcomes. In addition, therapists were all rated high on adherence and on competency, which could have led to a ceiling effect. Finally, we used a composite measure of competency, which attempted to measure this complex construct with a single item (the fact that moderate therapists were rated higher than both above- and below-average therapists in competency further underscores the latter two points). Perhaps a more sensitive measure would have yielded differences. More data are needed to clarify which of these possibilities best explains our findings regarding adherence and competency.

Consistent with the methodology used by Blatt et al. (1996) and Luborsky et al. (1997), therapists in the present study were rank ordered and categorized as above average and below average on the basis of the average therapeutic outcome scores of their patients. Although we used a similar idea of ranking, an important difference between our methods and Blatt et al. is that we found differences between individual therapists on a number of our key outcome measures in original ANCOVAs. Despite the fact that all but one therapist had caseloads that showed average improvement, differences between therapists existed. Therapists classified as above average differed significantly from the below average therapists in the average amount of change experienced by their patients, as rated on an outcome measure not used to rank the therapists, namely response status. C therapists were more likely to have patients that did not qualify as responders than were A therapists, both according to the independent measure of response and according to the more stringent criteria used to assess clinical significance. Thus, the therapists classified as above average were, indeed, more effective than the below-average therapists.

Through our analyses, we determined that differences between above-average and below-average therapists were attributable to above-average therapists' patients (a) being less likely to be nonresponders and (b) having greater improvement if they did respond to treatment. In fact, A therapists appeared to lower patients' anxiety sensitivity below the normal range, whereas C therapists' patients were clinically equivalent to nonpatients. However, A and C therapists did not differ on the number of dropouts in their caseload. Thus, the ability to cause more positive change in one's patients did not buffer patients' desire for dropping out from the study; neither did less effective therapists increase dropout rates. This result highlights the different aspects of therapists' effectiveness, including the ability to affect outcome and the ability to retain patients in treatment. Our data are consistent with the literature suggesting that experience with psychotherapy is one factor that is related to one's ability to affect outcome, whereas demographic factors and orientation are not likely to be relevant (Beutler et al., 1994).

There are other factors that are likely to have affected outcome as well. Behaviors that lead to the development and maintenance of a positive therapeutic relationship may distinguish A and C therapists (see Keijsers, Schaap, & Hoogduin, 2000, for a review). In addition, therapist technique and skill in CBT, in general, and in

treating panic, in particular, is worthy of further exploration, particularly because we used only a single global measure of competency. In addition, therapists' expectations, ability to get the patient to engage in homework consistently (Schmidt & Woolaway-Bickel, 2000), personality characteristics, and current life stressors are other potential candidates that could influence variation on outcome.

Therapist effect sizes were also calculated for each outcome measure, and consistent with Crits-Christoph et al. (1991), we found that therapist effects ranged from 1% for frequency of panic attacks to 17.9% for the ASI, with an 8% change on the PDSS (the main measure of panic disorder severity). These data support other calculations demonstrating that individual therapists make a modest contribution to outcome, even in a carefully monitored, manualized treatment for patients with panic disorder. These data lead to more questions about why some therapists differ from others in overall outcome and why the difference between therapists is greatest in a measure of anxiety sensitivity.

One unique aspect of this study is the examination of therapist rankings on each outcome measure used in the study. The use of multiple measures of panic disorder was an important contribution to addressing individual therapists' variability, which might have been missed otherwise. Table 3 shows that therapists were variably ranked on measures. Therapists were never ranked in exactly the same order on any two measures. Most of these measures are commonly used and have sound psychometric properties, making measurement error a less likely explanation for the variability in therapist ranking. In addition, all of these measures loaded onto a single factor, demonstrating that they are all related to an underlying construct that we argue is panic disorder severity. This makes measurement error even less likely. One possible explanation for therapist variability as indicated by different rankings on measures is that different measures might capture different elements of panic disorder, and variable rankings reflect therapists' greater and lesser ability to target different symptom domains.

There exists an intermediate group of therapists that varied substantially in ranking, depending on different outcome measures. Therapist 6 showed the greatest amount of change on the ASI (80% change for caseload) and average change on other measures (approximately 50% change), but this therapist also had the highest number (and percentage) of patients who dropped out of the study (7 out of 21). In contrast, Therapist 5 had the greatest change in panic disorder symptom severity according to the PDSS (80% change for caseload) but had the least amount of change in the ASI (10% change for caseload). Such information may have implications for understanding the mechanisms of change in CBT for panic.

Several limitations of these analyses exist. A relatively small number of clinicians were considered for these analyses and some treated only four patients. In addition, each of these clinicians was highly trained to a specific criterion so there was less variability among clinicians in terms of delivery of treatment. An ideal study of therapist effects would include all patients who are receiving psychotherapy alone, as opposed to the current study in which some patients were treated with combination therapy (i.e., imipramine or placebo). In addition, results should be considered with some caution because a large number of analyses were conducted with the relatively small therapist sample size of 14. Finally, cross-validation on a separate sample would demonstrate the va-

lidity of the groupings. Although this is usually impractical across studies (see Luborsky et al., 1997, for an exception), a study with a sufficient number of patients per therapist would allow a within-sample validation through ranking on a random half of the patients.

The results of this study raise a number of questions. What about therapist experience is important in influencing patient outcome? What specific factors differentiate the above- and below-average therapists? Could an analysis of the measures in which average therapists did more poorly help guide supervisors in determining where the therapists need more guidance? Are there any ways to help the below-average therapists improve outcomes to a level similar to the above-average therapists? Do therapists' differences in various outcome measures at posttreatment predict outcome during maintenance or follow-up stages of therapy? The answers to these questions have always been important. In the current age of behavioral health care and with the advent of quality assurance programs, ramifications of the answers to these questions are even more widespread. We need to further understand how therapists differ and whether they can be trained to be more effective. A study with sufficient therapist and patient sample size is needed, with the primary objective of further understanding therapist variables and their relationship to outcome in panic disorder.

References

Atkinson, D., & Schein, S. (1986). Similarity in counseling. *The Counseling Psychologist, 14,* 319–354.

Barlow, D. H. (1994). Psychological interventions in the era of managed competition. *Clinical Psychology: Science and Practice, 1,* 109–122.

Barlow, D. H. (1996). Health care policy, psychotherapy research, and the future of psychotherapy. *American Psychologist, 51,* 1050–1058.

Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: A randomized controlled trial. *Journal of the American Medical Association, 283,* 2529–2536.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1985). *Cognitive therapy of depression.* New York: Guilford Press.

Beck, D. F. (1988). *Counselor characteristics: How they affect outcomes.* Milwaukee, WI: Family Service America.

Beutler, L. E. (1997). The psychotherapist as a neglected variable in psychotherapy: An illustration by reference to the role of therapist experience and training. *Clinical Psychology: Science and Practice, 4,* 44–52.

Beutler, L. E., Machado, P. P., & Neufeldt, S. (1994). Therapist variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 229–269). New York: Wiley.

Blatt, S. J., Sanislow, C. A., Zuroff, D. C., & Pilkonis, P. A. (1996). Characteristics of effective therapists: Further analyses of data from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology, 64,* 1276–1284.

Bowman, D. G. (1993). Effects of therapist sex on the outcome of therapy. *Psychotherapy, 30,* 678–684.

Craske, M. G., Barlow, D. H., & Meadows, E. (2000). *Mastery of your anxiety and panic—3rd edition (MAP-3): Therapist guide for anxiety, panic, and agoraphobia.* San Antonio, TX: Graywind Publications/The Psychological Corporation.

Crits-Christoph, P., Baranackie, K., Kurcias, J. S., Beck, A. T., Carroll, K., Perry, K., Luborsky, L., McLellan, A. T., Woody, G. E., Thompson, L., Gallagher, D., & Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research, 1,* 81–91.

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for

the design and analysis of comparative studies of psychotherapy. *Journal of Consulting and Clinical Psychology, 59*, 20–26.

Dembo, R., Ikle, D. N., & Ciarlo, J. A. (1983). The influence of client–clinician demographic match on client treatment outcomes. *Journal of Psychiatric Treatment and Outcome, 5*, 45–53.

DiNardo, P. A., & Barlow, D. H. (1988). *Anxiety Disorders Interview Schedule—Revised (ADIS-R)*. Albany, NY: Graywind Publications.

Elkin, I. (1999). A major dilemma in psychotherapy outcome research: Disentangling therapists from therapies. *Clinical Psychology: Science and Practice, 6*, 10–32.

Elkin, I., Parloff, M. B., Hadley, S. W., & Autry, J. H. (1985). NIMH Treatment of Depression Collaborative Research Program: Background and research plan. *Archives of General Psychiatry, 42*, 305–316.

Fisher, E. H. (1989). Gender bias in therapy? An analysis of patient and therapist causal explanations. *Psychotherapy, 26*, 389–401.

Greenspan, M., & Kulish, N. M. (1985). Factors in premature termination in long-term psychotherapy. *Psychotherapy, 22*, 75–82.

Guy, W. (Ed.). (1976). *ECDEU assessment manual for psychopharmacology* (Publication No. ADM 76–338, pp. 217–222). Rockville, MD: U.S. Department of Health, Education, and Welfare.

Hafner, R. J., & Marks, I. M. (1976). Exposure in vivo of agoraphobics: Contributions of diazepam, group exposure, and anxiety evocation. *Psychological Medicine, 6*, 71–88.

Hamilton, M. (1959). The measurement of anxiety states by rating. *British Journal of Medical Psychology, 21*, 50–55.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23*, 56–62.

Huppert, J. D., Thorn, G. T., Hershberger, D. H., Hofmann, S., Barlow, D. H., Woods, S., Gorman, J., & Shear, M. K. (1997, November). *Therapist's experience after training and its nonsignificant relationship to outcome: Preliminary results from the Multicenter Collaborative Treatment Study of Panic Disorder.* Poster presented at the annual conference for the Association for the Advancement of Behavior Therapy, Miami, FL.

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.

Kazdin, A. (Ed.). (1997). The therapist as a neglected variable in psychotherapy research [Special section]. *Clinical Psychology, Science and Practice, 4*, 40–89.

Keijsers, G. P. J., Schaap, C. P. D. R., & Hoogduin, C. A. L. (2000). The impact of interpersonal patient and therapist behavior on outcome in cognitive–behavioral therapy. *Behavior Modification, 24*, 264–297.

Kendall, P. C., & Chambless, D. L. (Eds.). (1998). Empirically supported psychological therapies [Special section]. *Journal of Consulting and Clinical Psychology, 66*, 1–167.

Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 285–299.

Lafferty, P., Beutler, L. E., & Crago, M. (1989). Differences between more and less effective psychotherapists: A study of select therapist variables. *Journal of Consulting and Clinical Psychology, 57*, 76–80.

Luborsky, L., Auerbach, A. H., Chandler, M., Cohen, J., & Bachrach, H. (1971). Factors influencing the outcome of psychotherapy: A review of quantitative research. *Psychological Bulletin, 75*, 145–185.

Luborsky, L., Crits-Christoph, P., McLellan, T., Woody, G., Piper, W., Imber, S., & Liberman, B. (1986). Do therapists vary much in their success? Findings from four outcome studies. *American Journal of Orthopsychiatry, 56*, 501–512.

Luborsky, L., McLellan, A. T., Diguer, L., Woody, G., & Seligman, D. A. (1997). The psychotherapist matters: Comparison of outcomes across twenty-two therapists and seven patient samples. *Clinical Psychology: Science and Practice, 4*, 53–63.

Luborsky, L., McLellan, A. T., Woody, G. E., O'Brien, C. P., & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry, 42*, 602–611.

Martindale, C. (1978). The therapist-as-fixed-effect fallacy in psychotherapy research. *Journal of Consulting and Clinical Psychology, 46*, 1526–1530.

Nathan, P. E., & Gorman, J. M. (Eds.). (1998). *A guide to treatments that work.* New York: Oxford University Press.

Persons, R., Persons, M., & Newmark, I. (1974). Perceived helpful therapists' characteristics, client improvements, and sex of therapist and client. *Psychotherapy: Theory, Research, and Practice, 11*, 63–65.

Peterson, R. A., & Reiss, S. (1992). *Anxiety Sensitivity Index revised manual.* Worthington, OH: International Diagnostic Systems.

Ricks, D. F. (1974). Supershrink: Methods of a therapist judged successful on the basis of adult outcome of adolescent patients. In D. Ricks, M. Roff, & A. Thomas (Eds.), *Life history research in psychopathology* (Vol. 3, pp. 288–308). Minneapolis, MN: Minneapolis University Press.

Schmidt, N. B., & Woolaway-Bickel, K. (2000). The effects of treatment compliance on outcome in cognitive–behavioral therapy for panic disorder: Quality versus quantity. *Journal of Consulting and Clinical Psychology, 68*, 13–18.

Sexton, T. L., & Whiston, S. C. (1991). A review of the empirical basis for counseling: Implications for practice and training. *Counselor Education and Supervision, 30*, 330–354.

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmsted, M., Vallis, T. M., Dobson, K. S., Lowery, A., Sotsky, S. M., Watkins, J. T., & Imber, S. D. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67*, 837–846.

Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomskas, D. E., Woods, S. W., Gorman, J. M., & Papp, L. A. (1997). Multicenter collaborative Panic Disorder Severity Scale. *American Journal of Psychiatry, 154*, 1571–1575.

Stein, D. M., & Lambert, M. J. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced? *Journal of Consulting and Clinical Psychology, 63*, 182–196.

Strupp, H. H., & Anderson, T. (1997). On the limitations of therapy manuals. *Clinical Psychology: Science and Practice, 4*, 76–82.

Taylor, S. (Ed.). (1999). *Anxiety sensitivity.* Mahwah, NJ: Erlbaum.

Thompson, L. W., Gallagher, D., & Breckenridge, J. (1987). Comparative effectiveness of psychotherapies for depressed elders. *Journal of Consulting and Clinical Psychology, 55*, 385–390.

Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin, 117*, 450–468.

Zlotnick, C., Elkin, I., & Shea, T. M. (1998). Does the gender of a patient or the gender of a therapist affect the treatment of patients with major depression? *Journal of Consulting and Clinical Psychology, 66*, 655–659.