Journal of Consulting and Clinical Psychology
2000, Vol. 68, No. 6, 947-956

0022-006X/00/$5.00 DOI: 10.1037//0022-006X.68.6.947

EMDR for Panic Disorder With Agoraphobia: Comparison With Waiting List and Credible Attention-Placebo Control Conditions

Alan J. Goldstein and Edwin de Beurs
University of North Carolina at Chapel Hill
and Agoraphobia and Anxiety Treatment Center

Dianne L. Chambless and Kimberly A. Wilson
University of North Carolina at Chapel Hill

In a randomized controlled trial, eye movement desensitization and reprocessing (EMDR) for panic disorder with agoraphobia (PDA) was compared with both waiting list and credible attention-placebo control groups. EMDR was significantly better than waiting list for some outcome measures (questionnaire, diary, and interview measures of severity of anxiety, panic disorder, and agoraphobia) but not for others (panic attack frequency and anxious cognitions). However, low power and, for panic frequency, floor effects may account for these negative results. Differences between EMDR and the attention-placebo control condition were not statistically significant on any measure, and, in this case, the effect sizes were generally small (η^2 = .00–.06), suggesting the poor results for EMDR were not due to lack of power. Because there are established effective treatments such as cognitive–behavior therapy for PDA, these data, unless contradicted by future research, indicate EMDR should not be the first-line treatment for this disorder.

Although a relatively new and still controversial treatment, eye movement desensitization and reprocessing (EMDR; Shapiro, 1995) has spawned research concerning its efficacy and active elements, the bulk of this on posttraumatic stress disorder (PTSD). Much of this research is methodologically flawed (see Feske, 1998); however, there are a few well-designed studies documenting the superiority for PTSD of EMDR to a waiting list control group (Rothbaum, 1997; Wilson, Becker, & Tinker, 1995) or, in one case, to supportive listening (Scheck, Schaeffer, & Gillette, 1998). After a rigorous review of this literature, several authors (Chambless et al., 1998; DeRubeis & Crits-Christoph, 1998; Feske, 1998) concluded that EMDR appears to be beneficial for civilian PTSD but that its efficacy for combat-related PTSD remains to be substantiated in well-designed research. In a meta-analysis of PTSD studies, van Etten and Taylor (1998) found that, on most posttest measures of PTSD, EMDR was comparable to behavior therapy (a grouping of two effective treatments: exposure or stress inoculation training) and selective serotonin reuptake inhibitors (SSRIs) in efficacy and was superior to control conditions. Follow-up effect sizes for EMDR and behavior therapy continued to be equivalent. These results must be taken with caution because in no study included was EMDR directly compared with behavior therapy or SSRIs. Nonetheless, it is reasonable to conclude that EMDR offers some benefit for some forms of PTSD, at least when the trauma is not related to combat, even though the mechanisms for this efficacy remain in dispute (Steketee & Goldstein, 1994).

Shapiro (1995), who developed EMDR, has advocated its use for a wide variety of disorders other than PTSD, including other anxiety disorders. Uncontrolled case studies suggest potential benefits for panic and phobias (see, e.g., Goldstein & Feske, 1994; Marquis, 1991), but controlled research is sparse. One research group in the Netherlands is responsible for the bulk of EMDR for phobia research conducted with randomized controlled designs and sound measurement.[1] In two studies, Muris and Merckelbach and their colleagues compared EMDR with a waiting list (Muris & Merckelbach, 1997) or attention-placebo control group (Muris, Merckelbach, Holdrinet, & Sijsenaar, 1998) for treatment of spider phobia in adults or children, respectively. Treatments were limited to one 1- or 2.5-hr session. On behavioral avoidance tests, EMDR was not superior to the control conditions, but on a self-report

Alan J. Goldstein and Edwin de Beurs, Department of Psychology, University of North Carolina at Chapel Hill, and Agoraphobia and Anxiety Treatment Center, Bala Cynwyd, Pennsylvania; Dianne L. Chambless and Kimberly A. Wilson, Department of Psychology, University of North Carolina at Chapel Hill.

Edwin de Beurs is now at Vakgroep Psychiatrie, Vrije Universiteit, Amsterdam, the Netherlands.

This research was supported by National Institutes of Health Grant R21-MH49851. We thank project therapists Elisa Bolton, Kristina Coop Gordon, Barton Mann, Georgia Sloan, Taka Suzuki, Brigit van Widenfelt, and Mary Ellen Voegler, and research assistants Marnita Floyd and Maria Belecanech Beazley for their contributions to this research. We are especially grateful to Francine Shapiro for her generosity in providing training for the therapists and reviewing the treatment manual and to Ulrike Feske for invaluable assistance in the initiation of this project.

Correspondence concerning this article should be addressed to Alan J. Goldstein, Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3270. Electronic mail may be sent to alangold@email.unc.edu.

[1] Bates, McGlynn, Montgomery, and Mattke (1996) also compared one session of an EMDR analogue to an assessment-only control group. However, project personnel were not trained in EMDR, and significant aspects of the procedure appear to have been omitted or incorrectly applied (de Jongh, Ten Broeke, & Renssen, 1999; Lipke, 1997; McGlynn, 1997), making the relevance of the study's findings to Shapiro's (1995) EMDR uncertain.

measure of phobia (used only in the study with children), EMDR participants were significantly more improved than the control group children. Comparing EMDR with an assessment-only group for public-speaking anxiety (participants were not required to meet criteria for social phobia), Foley and Spates (1995) found one to two sessions of EMDR to be significantly more effective for two self-report outcome measures but not for heart rate or observable anxiety during a speech. However, the sample size was quite small (8–10 per group), and, for all measures except heart rate, the pattern of the data indicated more change in the EMDR than in the assessment-only group.

Testing EMDR against an active treatment, Muris, Merckelbach, and colleagues (Muris & Merckelbach, 1997; Muris et al., 1998; Muris, Merckelbach, van Haaften, & Mayer, 1997) compared one session of EMDR with one session of imaginal or in vivo exposure for spider phobia. Sessions were 1, 1.5, or 2.5 hr in duration, depending on the study. In general, in vivo exposure proved superior to EMDR on self-report measures and the behavioral avoidance test, with results sometimes significantly different and other times showing trends favoring exposure. The effects of one session of EMDR versus one session of imaginal flooding were not significantly different. Overall, these studies provide little support for EMDR's efficacy as a treatment of animal phobia. As the authors acknowledged, there are a number of methodological limitations to this research, including the questionable external validity of 1-hr treatment, failure to assess treatment integrity, and, in two studies, confounding of therapists and therapists' experience with treatment condition. Moreover, samples sizes were quite small (8 or 9 participants per group), yielding adequate power to detect only very large between-groups effects (e.g., Cohen's d of ≥ 1.07).

In a somewhat larger, more clinically representative study, Feske and Goldstein (1997) compared six sessions of EMDR (n = 15) with a waiting list control condition (n = 12) for panic disorder (almost always with agoraphobia). Clients were randomly assigned to treatment condition, and therapists were crossed with condition. Therapists were trained in EMDR by Shapiro, and their adherence to a treatment manual approved by Shapiro was verified by independent integrity raters. Assessment included reliable and valid self-report questionnaires of phobia, anxiety, and panic, and a daily panic and anxiety diary. EMDR was significantly superior to the waiting list on all measures. Despite the greater difficulty in successfully treating agoraphobic clients versus those with specific phobias (Chambless & Woody, 1990), this trial yielded more supportive findings for the efficacy of EMDR than the prior research on spider phobia. Possible reasons include the therapists' more extensive training in EMDR, the greater length of treatment, and increased statistical power.

Although Feske and Goldstein (1997) obtained positive results for EMDR's efficacy for panic disorder with agoraphobia (PDA), in light of the conflicting findings with earlier results for treatment of phobia, these effects require replication. Moreover, Feske and Goldstein's study is only a beginning step, in that a waiting list control condition fails to control for important nonspecific variables such as expectancy and therapists' attention. Although Muris et al. (1998) included an attention-placebo control group, they did not verify that this condition generated equal expectancy to EMDR. Accordingly, the purposes of the present study were twofold: (a) to conduct a replication of Feske and Goldstein's comparison of EMDR to a waiting list control group for PDA and (b) to contrast EMDR with a credible attention-placebo control group.

Method

Design

Participants were initially randomly assigned to one of three groups: waiting list (n = 14), EMDR (n = 18), or an attention-placebo condition (n = 13) involving the same amount of therapist contact as EMDR. Once the waiting list period ended, all those assigned to waiting list were randomized to EMDR (n = 6) or attention-placebo (n = 7). Follow-up of these two treatment conditions was accomplished 1 month after posttest; clients received no additional treatment during this interval. Therapists were crossed with treatment condition and were randomly assigned clients within scheduling constraints. Dropouts were replaced with the next participant to enter the study.

Participants

Inclusion and exclusion criteria. Participants were 46 outpatients applying for treatment at the Agoraphobia and Anxiety Treatment Center in Bala Cynwyd, PA (n = 12), or the Anxiety Treatment Center of the University of North Carolina at Chapel Hill (n = 34). All met *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.: *DSM–IV;* American Psychiatric Association, 1994) criteria for PDA of at least 1 year's duration. The agoraphobic avoidance had to have been of at least moderate severity for the prior 6 months.

Exclusion criteria included age less than 18 or greater than 65 and being in therapy elsewhere if not willing to suspend that treatment until the end of the study. Potential participants on dosages of alprazolam in excess of 1.5 mg daily (or similar dosages for other benzodiazepines) were excluded,[2] as were those who had been taking antidepressant or antianxiety medication for less than 6 months or who had changed their medication within the last 12 weeks. These exclusions were accomplished during telephone screening, and the number of potential participants excluded for these reasons was not recorded. Participants excluded on the basis of recent medication changes were eligible for reconsideration once medications were stabilized in appropriate limits. Potential participants were also excluded if they had comorbid diagnoses of thought disorder, major depression (n = 5), bipolar disorder, or substance dependence (n = 1); if another anxiety disorder was more severe than the PDA (n = 3); or if they met full criteria for any of the following Axis II disorders: paranoid (n = 1), schizoid, schizotypal, antisocial, or borderline (n = 3).

Sample characteristics. Participants' mean age was 38.16 years (range = 22–63). Thirty-seven were female. Two were African American, and 1 was Asian American; the remainder were European American. Thirty-eight had attended at least some college. Median and modal household socioeconomic status as indicated by occupation was 6 (e.g., technician, small business owner) on the Hollingshead (1975) scale. Mean duration of panic disorder was 12.56 years (range = 1–30). Twenty-one participants were taking psychotropic medication. Twenty participants had at least one comorbid Axis I diagnosis: specific phobia (7), generalized anxiety disorder (6), social phobia (5), or obsessive–compulsive disorder (2). Of these, 5 had more than one Axis I comorbid condition. Three participants met criteria for obsessive–compulsive personality disorder, and 4 for avoidant personality disorder.

[2] This criterion was set because there is some possibility that high doses of benzodiazepines and sedatives interfere with fear reduction (see, e.g., Chambless, Foa, Groves, & Goldstein, 1979; Marks et al., 1993).

Measures

The assessment battery of self-report questionnaires and diary and interview measures tapping comorbidity, panic, agoraphobia, and related constructs was designed in accordance with the recommendations from the consensus conference on assessment of panic disorder (Shear & Maser, 1994).

Structured Clinical Interview for DSM–IV (SCID). All diagnoses were based on the Structured Clinical Interview for *DSM–IV* (SCID; Axis I: First, Spitzer, Gibbon, & Williams, 1995; Axis II: First, Spitzer, Gibbon, Williams, & Benjamin, 1994), which was administered by clinical psychology doctoral students who were trained according to instructions from the SCID manual. Twenty-five audiotaped SCIDs were independently rated for reliability by a second rater. Interrater agreement for the diagnosis of PDA was 100%.

Self-report and interview outcome measures. Self-report instruments were completed 1 week prior to the onset of treatment (and waiting list if applicable), 1 week following the final treatment session, and 5–6 weeks following the final treatment session. These measures all have positive data on their reliability and validity. The Agoraphobic Cognitions Questionnaire (ACQ) and Body Sensations Questionnaire (BSQ; Chambless, Caputo, Bright, & Gallagher, 1984) were administered to assess fear of fear. The seven panic-related items of the Brief Body Sensations Interpretation Questionnaire (BBSIQ; Clark et al., 1997) were included to measure catastrophic interpretations of somatic symptoms typical of panic. The Panic Appraisal Inventory (PAI; Telch, Brouillard, Telch, Agras, & Taylor, 1989) was used to assess cognitions related to panic: feared consequences, anticipation, and appraisal of coping resources. The Mobility Inventory (Chambless, Caputo, Jasin, Gracely, & Williams, 1985) was employed to measure agoraphobic avoidance. To assess depressive symptoms and general anxiety, we included the Beck Depression Inventory (BDI; Beck & Steer, 1987) and the Beck Anxiety Inventory (BAI; Beck & Steer, 1990), respectively, in the battery. To serve as measures of global functioning, the Brief Symptom Inventory (BSI; Derogatis, 1992), the Social Adjustment Scale–Self-Report (SAS–SR; Weissman & Bothwell, 1976), and the Distress Questionnaire (adapted from Margraf & Schneider, 1990) were used. The Distress Questionnaire comprises four items on which patients rate the extent to which their anxiety interferes with their functioning on work or education, leisure and social activities, family relations, and household chores.

In addition to these self-report instruments, the Panic Disorder Symptom Severity Interview (PDSS; Shear et al., 1997) was conducted at each assessment point. A subset of the PDSS interviews was audiotaped and independently rated for reliability (n = 38). Reliability between interviewers was high, r = .91. Raters were not blind to group assignment.

Self-monitoring forms. For 2 weeks prior to and after treatment or waiting list, as well as throughout the course of treatment or waiting period, participants completed anxiety forms every morning and evening and at the close of each week. On these forms, participants were asked to rate on 10-point Likert-type scales their daily expectancy of having a panic attack, daily highest and average anxiety, daily number of panic attacks experienced, and, at the end of the week, expectancy of having a panic attack during the following week. Test–retest reliability for these ratings (aggregated into weekly blocks) ranged from .77 to .91 (de Beurs, Chambless, & Goldstein, 1997).

Treatment integrity. Immediately following the first session of treatment, participants completed the Treatment Expectancy Scale (Borkovec & Nau, 1972), rating the credibility of the treatment that had just been described to them in detail. After the next session, they also completed the Therapist Rating Scale (Williams & Chambless, 1990), in which they assessed the degree to which their therapist was caring, confident, accepting, challenging, explicit, and willing to be known. These subscales are internally consistent (α = .71–.94) and have proved predictive of treatment outcome for agoraphobic clients in behavior therapy (Williams & Chambless, 1990).

To ensure that therapists adhered to the treatment protocol, all sessions were audio- or videotaped and reviewed by Alan J. Goldstein prior to supervision meetings, which were held weekly to discuss clinical issues and proper provision of treatment. Two of the authors, Dianne L. Chambless and Kimberly A. Wilson, and their trained research assistants followed detailed integrity checklists that assessed adherence to treatment protocol, presence of therapist support and reinforcement, and protocol violations, which included introducing other treatments into the session. Adherence checks were conducted on 31% (n = 80) of all sessions. Of these, 33 were independently rated by additional coders to assess reliability. Average percent agreement was 95% for the integrity items identified a priori to be most important. The adherence monitoring team was not otherwise involved in participants' treatment and was unaware of participants' treatment outcome. Decisions to reject cases on the basis of poor therapist adherence rested solely with this independent team.

Treatment

Both treatments consisted of six 90-min sessions held over an average of 4 weeks and conducted according to detailed treatment manuals.[3] In both conditions, the aim of the first session was to gather information about the symptoms, history, and course of the disorder, as well as detailed descriptions of important memories related to panic attacks (e.g., the first attack, the worst attack). These descriptions later served as initial images for treatment sessions in both groups. Therapists also devoted substantial attention to providing participants with equally plausible rationales for both treatments. Indeed, on average, clients in both groups rated the rationale credible and endorsed positive expectancy that the treatment would be helpful (Ms = 6.48 and 7.70 for EMDR and attention-placebo, respectively, indicating greater than moderate credibility/expectancy). Clients in the attention-placebo group actually rated their forthcoming treatment significantly more positively than did clients in the EMDR group (Z = -2.34, $p < .05$).

Throughout treatment, therapists in both conditions were prohibited from using interventions outside the realm of the protocol such as anxiety management training, cognitive restructuring, in vivo exposure, and exploration of intrapsychic issues.

EMDR. EMDR was delivered according to a manual that had been developed in two previous studies on the treatment's efficacy (Feske & Goldstein, 1997; Goldstein & Feske, 1994) and reviewed and approved by Francine Shapiro, who developed EMDR. In the EMDR condition, therapists described EMDR as a method for facilitating reprocessing of traumatically induced fear networks. They further indicated that, although there had been prior success with EMDR for PDA, it was still an experimental treatment that might or might not be helpful for the client's condition.

Therapists initiated the EMDR process by having patients describe a pertinent, previously selected memory that continued to provoke significantly anxiety (e.g., their worst or first panic attack, or a body sensation of which they were especially fearful, such as heart palpitations). This step was followed by a set of eye movements, in which patients' eyes followed therapists' fingers moving side to side for approximately 20 s. Clients were then asked to indicate "what comes up," followed by another set of eye movements. This process continued until the original memory no longer elicited substantial anxiety (as determined by ratings of 0 or 1 on a 0–10 client-report anxiety scale) or until time ran out in the session. Once the scene had been desensitized, previously identified positive cognitions (e.g., "I'll be OK even if I panic") were paired with the original scene and a set of eye movements. Then clients provided a validity of cognition rating to indicate the extent to which they felt that the statement was true. Through-

[3] Copies of the manuals are available for the cost of duplication and postage from Alan J. Goldstein.

out the sessions, therapists were to prompt clients in a specified fashion if clients were not making progress as defined by decision rules in the manual. Otherwise, therapists were to allow the participants to continue the process with minimal interference or direction. At the close of each session, therapists debriefed participants by briefly discussing the themes that had emerged.

Adherence coders identified major violations (e.g., failure to engage in at least 45 min of EMDR during ≥2 sessions) during the treatment of 2 participants in the EMDR condition, both treated by the same therapist. These cases were omitted from further analyses, including the intention-to-treat analyses.

Association and relaxation therapy (ART). ART, the attention-placebo treatment, included a combination of two relatively inert treatment procedures: 30–45 min of progressive muscle relaxation training and 45–60 min of association therapy. Relaxation was conducted according to procedures described by Bernstein and Borkovec (1973); training in application of relaxation to anxiety-provoking situations was explicitly prohibited. Progressive relaxation without the anxiety management component has been shown to have little benefit for PDA clients (Chambless, Foa, Groves, & Goldstein, 1979; Öst, 1988). The associative therapy component was based on a treatment designed by Gelder et al. (1973) and used by these authors and others (Butler, Cullington, Munby, Amies, & Gelder, 1984; Taylor et al., 1997) as an attention-placebo condition. In all studies, associative therapy has proved to be significantly inferior to active cognitive–behavioral treatments for anxiety disorders.

Therapists described ART as an experimental approach that might work by helping clients to understand the meaning of their panic through a process of free association to memories of panic attacks. Clients were further told that they would be taught to relax to enhance the flow of associations. Sessions began with relaxation training, after which the remainder of each session was spent in association. As was the case for EMDR, the starting point for association was the most frightening panic-related image identified in the initial session. Therapists asked clients to begin by describing the scene in detail and then to close their eyes (or look at a fixed place) and simply speak out loud as they allowed themselves to associate to this image. Therapists remained quiet and fairly passive unless the client reported difficulty continuing or initiating associations. In this case, the therapist would provide standardized prompts, such as "Let come up whatever comes up." Therapists were to instruct clients not to engage in conversation but to continue associating until the end of the session, when they very briefly discussed the day's themes. No significant protocol violations were observed in this condition.

Therapists. All therapy was conducted by doctoral or postdoctoral students in clinical psychology ($n = 4$) or by masters- or doctoral-level clinicians ($n = 4$) with 5–11 years of postdegree clinical experience. All therapists were female save one, and all had received Level II (the highest level) training from either Francine Shapiro or one of her appointees. Therapists treated from 3–15 cases, depending on length of employment. There was no significant difference in the number of clients per condition therapists treated, $\chi^2(6, N = 45) = 1.40$, $p > .96$.

Results

Preliminary Analyses

Attrition. Of the 46 participants who entered the study, 4 dropped out prior to the completion of treatment. One dropped out during the waiting list period before she provided posttest data or received her treatment condition assignment. Three participants (one of whom had previously been in the waiting list condition) dropped out or were terminated during EMDR: one because of a marital crisis, another because of deterioration, and a third for repeated cancellations of appointments. Fisher's exact tests indicated that attrition was not significantly different across groups (EMDR vs. attention-placebo $p = .242$; EMDR vs. waiting list $p = 1.00$). Of the 42 participants who completed treatment, 37 provided follow-up data. Of those who dropped from follow-up after EMDR, one required medical attention for an unrelated condition, one terminated because of increased distress during treatment, and a third refused assessment without explanation. Of those who dropped from attention-placebo, one dropped because of his disappointment with treatment and the other without explanation.

Inclusion of cases at each assessment point. Included in the comparison of EMDR to waiting list were 27 participants (14 EMDR, 13 waiting list). After the month-long waiting list period, clients in that condition were randomly reassigned to EMDR ($n = 6$) or ART ($n = 7$), thus adding to those assigned immediately to those groups to yield a sample size of 20 each for EMDR versus ART comparisons.

Power analyses. For the comparison of EMDR to waiting list, based on alpha of .05 and 27 participants, we had 80% power to detect only large effect sizes. Cast as η^2, this was equivalent to an effect size of ≥.23. For the comparison of EMDR to ART at posttest, based on alpha of .05 and 40 participants, we had 80% power to detect large effect sizes, defined as $\eta^2 \geq .168$; for follow-up analyses with 35 participants, the comparable η^2 was .187.

Composites. Composite variables were created to serve as reliable representations of important constructs and to reduce the number of statistical tests to be performed. Principal components factor analyses were conducted on anxiety-related self-report measures and interview, global functioning measures, and diary variables. Scree plots were used to determine the number of factors. Composites were then computed by adding the variables (standardized according to pretest variance) that loaded on a given factor.

Factor analysis of the anxiety-related interview and self-report measures yielded a two-factor solution that accounted for 60% of the variance at pretest. These results were consistent with two conceptually distinct factors: measures of cognition and measures of panic/agoraphobia severity. Measures of cognition consisted of the ACQ, the BSQ, the BBSIQ, and the PAI's Consequences and Coping subscales. Measures of panic/agoraphobia severity consisted of the PDSS, the Mobility Inventory (Avoidance Alone), the BDI, the BAI, and the PAI's Anticipation subscale.

Items from the self-monitoring anxiety forms loaded onto one factor, accounting for 71% of the variance in pretest scores. Therefore, one composite variable was created from these items and labeled the diary variable. Scores on the three general functioning measures (the BSI Global Severity Index, the SAS–SR, and the Distress Questionnaire) also formed one factor, named global functioning, which accounted for 68% of the variance in pretest scores. Diary-assessed panic frequency was treated as a separate variable and defined as the average number of weekly panic attacks per week over the assessment interval (2 weeks each at pretest, posttest, and follow-up). Thus, the number of dependent variables was reduced to five: self-report and interview measures of severity of panic disorder/agoraphobia, diary measures, panic frequency, negative cognitions, and global functioning. All five variables were tested at follow-up, whereas, at posttest, tests of global functioning were omitted. Given the brevity of treatment

(3–4 weeks), we thought the pretest–posttest interval was inadequate to observe changes on measures of disability.

Covariates. Tests for differences between groups on attribute and outcome variables at pretest were conducted with Fisher's exact and Mann-Whitney tests. When differences were found, Spearman correlations were then calculated between those variables and residual gain scores to determine if the variables in question were related to outcome. In cases where these relationships tended to be significant ($p < .20$), these variables were included as covariates in subsequent posttest and follow-up comparisons.

The following demographic and diagnostic variables were examined: household mean education, presence of comorbid Axis I diagnoses, presence of Axis II diagnoses, total number of comorbid diagnoses, global assessment of functioning (GAF) from the SCID, duration of panic disorder, and whether the participant took medication. Differences were also examined on all dependent variables at pretest. For the EMDR versus ART comparisons, site effects, therapist ratings, and treatment expectancy ratings were also examined. The number of therapists relative to the sample size precluded statistical analysis of therapist effects on treatment outcome. No site effects were detected. In the analyses that follow, no trends for confounding variables were observed except as explicitly noted.

Major Analyses

Omnibus three-group analyses of variance (ANOVAs) were eschewed in favor of focused ANOVAs comparing EMDR with the waiting list condition and EMDR with ART.[4] Comparisons between the groups at posttest and follow-up were made with repeated measures ANOVAs and analyses of covariance (ANCOVAs). In light of the limited power, alpha was set at .05. For these parametric analyses, panic frequency was normalized through a log transformation because of the skewed distribution of the original variable.

EMDR versus waiting list. At pretest, participants in the EMDR condition ($n = 14$) reported more panic attacks than those in the waiting list condition ($n = 13$; $Z = -1.71$, $p < .10$). This pretest variable tended to predict outcome on cognitive factor ($r = -.30$, $p < .20$), therefore warranting its inclusion in subsequent analyses as a covariate.

Repeated measures ANOVAs or ANCOVAs showed that EMDR was significantly superior to waiting list on panic/agoraphobia severity and on the diary factor ($ps < .05$) but not on the cognitive factor, controlling for pretest panic frequency ($p > .10$). Furthermore, EMDR did not show greater improvement over waiting list on number of panic attacks ($p > .10$), as indicated by ANCOVA on posttest scores, controlling for pretest panic frequency. See Table 1 for group means and standard deviations on individual measures and Table 2 for within-group effect sizes for each condition. Table 3 summarizes between-groups differences. The between-groups effect sizes are medium to very large, always favoring EMDR.

EMDR versus ART. Because waiting list clients were randomly reassigned to ART or EMDR at the end of the waiting period, preliminary ANOVAs were conducted to test whether post-waiting-list scores on dependent measures for control group clients were significantly different than pretest scores for those who had entered EMDR or ART straightaway. In no case did the F tests approach significance (all $ps > .13$), and no medium or larger effects sizes (e.g., $\eta^2 > .058$) were obtained. Accordingly, initial group membership was not considered further as a factor in comparison of EMDR versus ART.

Participants in the ART condition ($n = 20$) had a longer duration of panic disorder than did those assigned to EMDR ($n = 20$). Because duration predicted better outcome on the diary factor ($r = -.28$, $p < .01$), it was included as a covariate in that analysis. Because treatment expectancy tended to predict better outcome on panic/agoraphobia severity ($r = -.27$, $p = .10$) and expectancy was significantly higher for the ART group, this variable was included as a covariate for that comparison. No differences between treatment groups emerged ($ps > .10$). See Table 4 for group means and standard deviations on individual measures and Table 5 for within-group effect sizes for each treatment condition. Table 6 summarizes between-groups findings.

Follow-up. When tested before treatment, participants in the ART condition who provided follow-up data ($n = 18$) had a longer duration of panic disorder ($Z = -1.39$, $p < .20$), lower GAF ($Z = -1.31$, $p < .20$), and lower scores on the functioning composite ($Z = -1.91$, $p < .10$) than those providing data in the EMDR condition ($n = 17$). Because both longer duration of panic disorder and lower GAF tended to predict outcome on functioning ($rs = .34$ and $-.28$, respectively, $ps < .20$), these variables were entered as covariates for between-groups comparisons on functioning outcome. Duration also tended to predict panic/agoraphobia severity at follow-up ($r = .24$, $p < .20$) and was thus treated as a covariate for that analysis. In addition, participants in ART included at this time point rated their treatment expectancy higher than those in EMDR ($Ms = 6.65$ and 7.68, respectively; $Z = -2.04$, $p < .05$). Treatment expectancy tended to predict follow-up panic frequency ($r = -.35$, $p < .10$) and, therefore, served as a covariate for that analysis. No differences emerged on any outcome variables at follow-up ($ps > .10$). Descriptive data are reported in Table 4 and between-groups comparisons in Table 6.

Intention-to-treat analyses. Intention-to-treat analyses were conducted at each assessment period by repeating ANOVAs and ANCOVAs with pretest scores carried forward to serve as posttest or follow-up scores for those who failed to provide posttest data or who dropped out before the conclusion of treatment or before the follow-up assessment. The findings of the EMDR versus waiting list and EMDR versus ART comparisons were unchanged.

Clinical significance. Clinical significance of EMDR's effects was tested in a series of comparisons with normal sample data provided by Gillis, Haaga, and Ford (1995), Bibb (1988), and Clark et al. (1997). These comparisons were limited to select outcome measures of central constructs for which normal sample data were available: the BAI, the Mobility Inventory for Agoraphobia Avoidance Alone scale, and the BBSIQ. Following procedures suggested by Kendall, Marrs-Garcia, Nath, and Sheldrick (1999) and by Rogers, Howard, and Vessey (1993), we compared

[4] A between-groups ANOVA requires the assumption of independence of observations. Had we conducted three-group comparisons with waiting list participants included twice (once in the waiting list condition and once when they had completed the waiting period and had been rerandomized to EMDR or ART), this assumption would have been violated.

Table 1
Descriptive Data for EMDR and Waiting List Comparison: Factor Scores and Individual Dependent Measures

Measure	EMDR[a] (*SD*)	Waiting list[b] (*SD*)
Cognitive measures		
Pre	8.83 (2.96)	9.50 (3.94)
Post	6.75 (3.53)	8.79 (3.86)
Agoraphobic Cognitions Questionnaire (range = 1–5)		
Pre	2.09 (0.65)	2.32 (0.59)
Post	2.07 (0.73)	2.19 (0.51)
Body Sensations Questionnaire (range = 1–5)		
Pre	2.64 (0.68)	2.74 (0.86)
Post	2.41 (0.73)	2.67 (0.75)
Body Sensations Interpretation Questionnaire–Panic (range = 1–3)		
Pre	1.90 (0.72)	1.64 (0.60)
Post	1.63 (0.57)	1.56 (0.72)
Panic Appraisal Inventory–Coping (range = 0–100)		
Pre	35.98 (17.42)	30.84 (16.94)
Post	50.33 (22.22)	33.12 (14.02)
Panic Appraisal Inventory–Consequences (range = 0–100)		
Pre	51.54 (31.84)	62.46 (29.31)
Post	35.29 (31.02)	56.46 (34.16)
Panic and agoraphobia severity		
Pre	12.44 (3.68)	12.02 (4.11)
Post	9.43 (3.21)	12.04 (4.15)
Panic Disorder Severity Scale (Interview) (range = 0–4)		
Pre	2.13 (0.74)	1.96 (0.57)
Post	1.51 (0.68)	2.01 (0.70)
Mobility Inventory–Alone (range = 1–5)		
Pre	2.88 (0.80)	3.19 (0.96)
Post	2.42 (0.74)	3.07 (1.05)
Beck Depression Inventory (range = 0–63)		
Pre	15.00 (7.47)	12.74 (9.38)
Post	11.43 (6.95)	11.95 (7.31)
Beck Anxiety Inventory (range = 0–63)		
Pre	24.68 (10.62)	21.00 (10.66)
Post	15.57 (7.58)	22.00 (10.76)
Panic Appraisal Inventory–Anticipation (range = 0–100)		
Pre	46.81 (18.35)	47.47 (22.97)
Post	38.29 (20.85)	49.00 (21.12)
Diary		
Pre	8.88 (4.06)	9.23 (3.22)
Post	7.57 (4.27)	10.37 (3.50)
Average daily anxiety (range = 0–10)		
Pre	3.41 (1.42)	3.55 (1.56)
Post	2.84 (1.67)	4.38 (1.84)
Fear of panic (range = 0–10)		
Pre	4.39 (2.80)	4.92 (1.84)
Post	3.35 (2.12)	5.26 (2.14)
Panic attack expectancy–daily (range = 0–10)		
Pre	3.84 (2.86)	4.16 (1.58)
Post	3.49 (2.79)	4.52 (1.61)
Panic attack expectancy–weekly (range = 0–10)		
Pre	5.75 (2.82)	5.38 (1.69)
Post	5.18 (2.87)	5.92 (1.56)
Average panic attacks per week[c]		
Pre	1.45	0.50
Post	0.57	0.54
Log of panic frequency[c]		
Pre	0.47 (0.41)	0.21 (0.24)
Post	0.28 (0.37)	0.22 (0.24)

Note. Measures in bold type are factor scores. EMDR = eye movement desensitization and reprocessing.
[a] $n = 14$. [b] $n = 13$. [c] Medians (rather than means) and log transformations are provided for panic frequency because of the highly skewed distribution of this variable.

the posttest mean scores for the EMDR group with the normal sample data, conducting *t* tests for significant differences between these two groups, as well as *t* tests of the equivalence between these groups. We defined a difference of less than 1 *SD* between the normal and EMDR-treated PDA participants as equivalence. On the Mobility Inventory's Avoidance Alone scale, the PDA clients were significantly worse on both traditional and clinical equivalence *t* tests, $t_{trad}(103) = 30.21, p < .01$; $t_{equiv}(103) = 5.57, p < .01$. This also proved to be the case for the BBSIQ, $t_{trad}(38) = 3.57, p < .01$; $t_{equiv}(38) = 2.14, p < .05$. On the BAI, PDA clients were significantly worse than the normal group on the traditional *t* test, $t_{trad}(260) = 5.83, p < .01$, but were not significantly different on the test of clinical equivalence, $t_{equiv}(260) = 1.52, p > .05$. Finally, we examined the percentage of EMDR clients who were panic free in 2 weeks of daily monitoring. This figure rose from 25% at pretest to 45% at posttest, leaving the majority reporting at least one panic attack in a 2-week period at posttest.

Discussion

The first goal of the present study was a replication of Feske and Goldstein's (1997) comparison of EMDR with a waiting list control condition for PDA. The present results were somewhat less favorable than those of the earlier study. EMDR was significantly superior to the waiting list condition on only two of the four composite measures. Controlled comparisons demonstrated significant improvement for EMDR clients on self-report, diary, and interview measures of panic disorder severity, agoraphobia, and

Table 2
Pretest–Posttest Within-Group Effect Sizes (Cohen's d') for EMDR and Waiting List

Measure	EMDR (*n* = 14)	Waiting list (*n* = 13)
Cognitive measures	0.75	0.59
Panic and agoraphobic severity	0.72	0.01
Diary	0.41	−0.44[a]
Panic frequency	0.39	0.04

Note. EMDR = eye movement desensitization and reprocessing.
[a] This effect size reflects a deterioration of those in the waiting list on the diary factor.

Table 3
ANOVA or ANCOVA Comparisons Between EMDR and Waiting List on Outcome Variables

Measure	*F*	Partial η^2
Cognitive measures	2.69	.10
Panic and agoraphobic severity	9.91**	.28
Diary	5.80*	.20
Panic frequency	1.93	.08

Note. ANOVA = analysis of variance; ANCOVA = analysis of covariance; EMDR = eye movement desensitization and reprocessing.
* $p < .05$. ** $p < .01$.

anxiety, whereas no significant change was apparent on the cognitive measures or on panic attack frequency.

In part, this discrepancy may be attributed to low power, in that power was adequate for only very large effect sizes. In addition, differences in the study sample may have played a role. Unlike Feske and Goldstein's (1997) clients, the present participants were selected on the basis of the severity of their agoraphobia (requiring moderate–severe avoidance), with no threshold set for frequency of panic attacks in the pretreatment interval, so long as the client met diagnostic criteria for panic disorder. As a result, floor effects pose some difficulty for the panic frequency measure; 44% of the current sample recorded no panic attacks in the 2-week pretreatment monitoring interval. In contrast, Feske and Goldstein excluded clients who did not have at least one panic attack in this time period. Nonetheless, the posttest results for the EMDR group in which only 45% of clients were panic free are disappointing in comparison with the results of comparably brief cognitive therapy (Clark et al., 1999) in which 71% of clients reported no panic in the posttreatment monitoring period.

Floor effects are a less likely explanation for the lack of significant differences between EMDR and the waiting list condition on the cognitive measures factor. Despite the emphasis in EMDR on changing cognition through alteration of the fear network and pairing a realistic, positive thought with the desensitized frightening image, the treatment did not have a large effect on maladaptive thinking ($\eta^2 = .10$, a medium effect). For Feske and Goldstein's (1997) study, the comparable effect size is large ($\eta^2 = .14$) but is still greatly exceeded by results of comparisons of cognitive therapy with waiting list conditions for panic disorder. These comparisons have consistently yielded very large effects. For example, in their review of cognitive therapy versus waiting list and supportive therapy control groups, Chambless and Gillis (1993) obtained very large average effect sizes (recast here as η^2) ranging from .24 to .42 for measures of negative cognitions, anxiety, and panic attacks. Thus, although these authors did not specify which of these effect sizes pertained to negative cognitions, it was greater than or equal to .24, whereas it was only .10–.14 in our two studies of EMDR.

Overall, the present results, along with those of Feske and Goldstein (1997), appear to be more positive for EMDR than those reported by Muris and Merckelbach (1997), who provided the only other comparison with a waiting list control condition for phobia. Differences in attention to the external and internal validity of the research designs and power may account for the variance in findings. Muris and Merckelbach provided only 1 hr of EMDR, whereas clients in the other two studies had a more clinically representative treatment of five 1.5-hr EMDR sessions after an initial treatment planning session. In addition, in contrast to the other two studies, Muris and Merckelbach did not report using a treatment manual or integrity checks, and their therapist had only brief training in EMDR. Finally, given that these authors had only 8 participants per group, their power for detecting differences between groups was very low, only .31 to detect a large effect size. Indeed, cast as Cohen's *d*, their EMDR versus waiting list effect size for the behavioral avoidance test (the only outcome measure provided for this comparison) was .83, a large effect (see Cohen, 1988). For comparison, using the Mobility Inventory, we computed the EMDR versus waiting list effect size for phobic avoidance in the present study ($d = .86$) and that of Feske and Goldstein ($d = .63$). Effect sizes were similar to or even smaller than the effect obtained by Muris and Merckelbach, indicating that the apparent discrepancies are most likely a misleading consequence of differences in power among the three studies.

Our second goal was to contrast EMDR's efficacy with that of an attention-placebo control condition of demonstrated equal credibility/expectancy. To our knowledge, this is the first such study. In three prior studies (two with PTSD: Marcus, Marquis, & Sakai, 1997; Scheck et al., 1998; and one for spider phobia: Muris et al., 1998), investigators included attention-placebo or treatment-as-usual control groups but failed to report data on their credibility/expectancy. In the present study, EMDR clients fared no better than those in the attention-placebo group. Although power was admittedly low, the small sample size is unlikely to be at fault, in that effect sizes for group contrasts were mostly negligibly small. There were only two medium effect sizes. In one case, this favored EMDR (panic/agoraphobia severity at posttest) and, in the other, the attention-placebo group (global functioning at follow-up), suggesting a chance distribution.

It is unlikely that the unfavorable results for EMDR versus the attention-placebo are due to poor methodology. We used a randomized control group design with carefully diagnosed clients. Results were consistent across the multiple methods of measurement employed, including reliable and valid self-report, interview, and daily diary measures. Treatment was conducted according to a manual reviewed and approved by Shapiro. Therapists in this trial were trained in EMDR before seeing clients in the study and were supervised on a weekly basis throughout the study by Alan J. Goldstein, a therapist highly experienced in EMDR for panic and agoraphobia. Extensive integrity checks were conducted throughout the trial, with two EMDR cases being removed from the data set when the therapist was found to have deviated from protocol. Therapists treated clients in both conditions, thus avoiding confounding therapist factors with treatment. Therapists were not told that the attention-placebo group was intended as such; rather, they were told that this group was part of a dismantling study, testing the active components of EMDR. Nonetheless, therapists evidenced bias in favor of EMDR, commenting frequently that they felt guilty about doing so little for clients in the attention-placebo group. Still, EMDR failed to best the attention-placebo group.

One might argue that, although each of the components of the attention-placebo treatment has previously been shown to be a

Table 4
Descriptive Data for EMDR and Attention-Placebo Comparison at Pretest, Posttest, and 1-Month Follow-Up: Factor Scores and Individual Dependent Measures

Measure	EMDR[a] (*SD*)	ART[a] (*SD*)
Cognitive measures		
Pre	13.52 (2.95)	13.54 (3.35)
Post	13.46 (3.01)	13.31 (3.30)
Follow-up	13.07 (2.68)	12.85 (3.28)
Agoraphobic Cognitions Questionnaire		
Pre	2.16 (0.59)	2.30 (0.71)
Post	2.20 (0.68)	2.20 (0.57)
Follow-up	2.08 (0.60)	2.06 (0.59)
Body Sensations Questionnaire		
Pre	2.68 (0.63)	2.75 (0.93)
Post	2.51 (0.66)	2.42 (0.83)
Follow-up	2.49 (0.79)	2.34 (0.77)
Body Sensations Interpretation Questionnaire–Panic		
Pre	1.81 (0.72)	1.62 (0.59)
Post	1.60 (0.60)	1.54 (0.58)
Follow-up	1.44 (0.59)	1.47 (0.63)
Panic Appraisal Inventory–Coping		
Pre	34.72 (16.27)	38.76 (17.79)
Post	47.30 (21.52)	44.58 (16.83)
Follow-up	46.40 (24.67)	46.92 (13.23)
Panic Appraisal Inventory–Consequences		
Pre	57.33 (30.79)	53.11 (38.95)
Post	45.63 (32.81)	53.63 (34.80)
Follow-up	50.21 (36.94)	47.90 (36.66)
Panic and agoraphobia severity		
Pre	13.65 (3.68)	12.51 (3.57)
Post	10.93 (3.76)	10.50 (3.23)
Follow-up	10.70 (4.19)	10.21 (3.03)
Panic Disorder Severity Scale (*Interview*)		
Pre	2.19 (0.68)	2.12 (0.68)
Post	1.61 (0.67)	1.54 (0.52)
Follow-up	1.54 (0.67)	1.58 (0.50)
Mobility Inventory–Alone		
Pre	3.07 (0.92)	3.02 (0.91)
Post	2.73 (0.91)	3.01 (0.91)
Follow-up	2.71 (1.00)	2.78 (0.85)
Beck Depression Inventory		
Pre	15.10 (6.55)	11.31 (8.24)
Post	11.80 (6.23)	7.55 (5.83)
Follow-up	10.79 (7.38)	8.40 (7.56)
Beck Anxiety Inventory		
Pre	25.28 (9.23)	21.04 (11.96)
Post	17.56 (8.21)	14.91 (9.79)
Follow-up	17.95 (13.53)	16.30 (10.25)
Panic Appraisal Inventory–Anticipation		
Pre	48.25 (17.28)	46.35 (18.77)
Post	41.80 (21.88)	44.65 (19.36)
Follow-up	42.58 (19.32)	38.91 (16.06)
Diary		
Pre	10.04 (3.82)	8.69 (2.56)
Post	8.46 (3.97)	7.46 (2.54)
Follow-up	7.90 (4.09)	6.92 (2.31)
Average daily anxiety		
Pre	3.91 (1.44)	3.62 (1.51)
Post	3.31 (1.80)	3.17 (1.49)
Follow-up	3.02 (1.70)	2.82 (1.37)
Fear of panic		
Pre	4.93 (2.52)	3.69 (1.56)
Post	3.92 (2.08)	3.32 (1.43)
Follow-up	3.80 (2.27)	3.08 (1.38)
Diary (*continued*)		
Panic attack expectancy–Daily		
Pre	4.28 (2.52)	3.58 (1.56)
Post	3.84 (2.49)	3.25 (1.49)
Follow-up	3.59 (2.58)	3.09 (1.50)
Panic attack expectancy–weekly		
Pre	6.03 (2.45)	5.58 (1.45)
Post	5.05 (2.48)	4.67 (1.49)
Follow-up	4.68 (2.40)	4.15 (0.90)
Average panic attacks per week[c]		
Pre	1.20	0.70
Post	0.54	0.00
Follow-up	0.29	0.00
Log of panic frequency[c]		
Pre	0.40 (0.38)	0.26 (0.27)
Post	0.25 (0.35)	0.18 (0.19)
Follow-up	0.23 (0.29)	0.01 (0.12)
Global functioning[b]		
Pre	11.37 (2.06)	9.68 (2.34)
Follow-up	9.60 (3.05)	8.88 (2.66)
Brief Symptom Inventory (range = 0–4)		
Pre	1.16 (0.46)	0.77 (0.50)
Follow-up	0.88 (0.65)	0.70 (0.45)
Distress Questionnaire (range = 0–4)		
Pre	1.89 (0.73)	1.75 (0.74)
Follow-up	1.31 (0.81)	1.53 (0.81)
Social Adjustment Scale (range = 1–5)		
Pre	2.08 (0.34)	1.85 (0.29)
Follow-up	1.93 (0.37)	1.74 (0.39)

Note. Measures in bold type are factor scores. EMDR = eye movement desensitization and reprocessing; ART = association and relaxation therapy.
[a] $n = 20$.
[b] Global functioning assessed only at pretest and 1-month follow-up.
[c] Medians (rather than means) and log transformations are provided for panic frequency because of the highly skewed distribution of this variable.

relatively ineffective treatment, their combination might prove to be efficacious. For comparison purposes, we computed Cohen's *d* for Mavissakalian's (1987) data on response of agoraphobic clients to pill placebo according to measures of anxiety and agoraphobic

Table 5
Pretest–Posttest and Pretest–1-Month Follow-Up Within-Group Effect Sizes (Cohen's d') for EMDR and ART

Measure	EMDR	ART
Cognitive measures		
Post	0.03	0.11
Follow-up	0.18	0.29
Panic and agoraphobia severity		
Post	0.78	0.89
Follow-up	0.70	0.96
Diary		
Post	0.53	0.63
Follow-up	0.66	0.85
Panic frequency		
Post	0.23	0.47
Follow-up	0.65	0.65
Global functioning[a]		
Follow-up	0.91	0.47

Note. EMDR = eye movement desensitization and reprocessing; ART = association and relaxation therapy.
[a] Not assessed at posttest.

Table 6
ANOVA and ANCOVA Comparisons Between EMDR and ART on Outcome Variables: Pretest–Posttest and Pretest–1-Month Follow-Up

Time and variable	*F*	Partial η^2
Posttest		
Panic and agoraphobia severity	2.01	.06
Cognitive factor	0.04	.00
Diary	0.22	.01
Panic frequency	1.28	.03
Follow-up		
Panic and agoraphobia severity	0.41	.01
Cognitive factor	0.00	.00
Diary	1.49	.05
Panic frequency	0.72	.03
Global functioning	1.69	.06

Note. $p > .05$ for all comparisons. ANOVA = analysis of variance; ANCOVA = analysis of covariance; EMDR = eye movement desensitization and reprocessing; ART = association and relaxation therapy.

avoidance. Within-group effect sizes for a 2-week period were 0.64 and 0.65.[5] Comparable effect sizes for agoraphobic avoidance and anxiety for the attention-placebo condition in the present study were .01 and .56, respectively. Thus, response of our clients to attention-placebo was well within expected limits for a placebo control condition. Finally, we note that follow-up in our study was brief (1 month). Accordingly, one might question whether EMDR would have surpassed the attention-placebo group at longer term follow-up. We consider this to be unlikely. In our prior study (Feske & Goldstein, 1997), the follow-up was of 3 months' duration, and the differences between EMDR and the control condition shrank rather than grew during that period because of (nonsignificant) deterioration in the EMDR group.

In summary, although EMDR was superior to a waiting list control group in the present study (on two of four outcome measures), it was no better than a credible attention-placebo control condition. In contrast, the beneficial effects of cognitive therapy and exposure for panic disorder (see review by Chambless & Gillis, 1993) and agoraphobia (see meta-analysis by van Balkom et al., 1997) have been extensively documented in well-designed research studies. Moreover, Muris et al. (1997, 1998) found in vivo exposure to be superior to EMDR for spider phobia. In light of the availability of treatments with solid efficacy evidence, the results of this investigation do not support the use of EMDR for treatment of panic disorder with agoraphobia.

[5] Elsewhere in this article, we have used Cohen's *d'*, which is preferable for repeated measures data; however, Mavissakalian (1987) did not provide all the data necessary for a calculation of *d'*.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bates, L. W., McGlynn, F. D., Montgomery, R. W., & Mattke, T. (1996). Effects of eye-movement desensitization versus no treatment on repeated measures of fear of spiders. *Journal of Anxiety Disorders, 6,* 555–569.

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. (1990). *Manual for the Beck Anxiety Inventory.* San Antonio, TX: Psychological Corporation.

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

Bibb, J. L. (1988). *Parental bonding, pathological development, and fear of losing control among agoraphobics and normals.* Unpublished doctoral dissertation, American University, Washington, DC.

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3,* 257–260.

Butler, G., Cullington, A., Munby, M., Amies, P., & Gelder, M. (1984). Exposure and anxiety management in the treatment of social phobia. *Journal of Consulting and Clinical Psychology, 52,* 642–650.

Chambless, D. L., Baker, M., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., Daiuto, A., DeRubeis, R., Detweiler, J., Haaga, D. A. F., Bennett Johnson, S., McCurry, S., Mueser, K. T., Pope, K. S., Sanderson, W. C., Shoham, V., Stickle, T., Williams, D. A., & Woody, S. R. (1998). Update on empirically validated therapies, II. *Clinical Psychologist, 51,* 3–16.

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

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E., & Williams, C. (1985). The Mobility Inventory for Agoraphobia. *Behaviour Research and Therapy, 23,* 35–44.

Chambless, D. L., Foa, E. B., Groves, G. A., & Goldstein, A. J. (1979). Flooding with Brevital in the treatment of agoraphobia: Countereffective? *Behaviour Research and Therapy, 17,* 243–251.

Chambless, D. L., & Gillis, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology, 61,* 248–260.

Chambless, D. L., & Woody, S. R. (1990). Are agoraphobics really harder to treat? A comparison of agoraphobics' and simple phobics' responses to in vivo exposure. *Behaviour Research and Therapy, 28,* 305–312.

Clark, D. M., Salkovskis, P. M., Hackmann, A., Wells, A., Ludgate, J., & Gelder, M. (1999). Brief cognitive therapy for panic disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 67,* 583–589.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

de Beurs, E., Chambless, D. L., & Goldstein, A. J. (1997). Measurement of panic disorder by a modified panic diary. *Depression and Anxiety, 6,* 133–139.

de Jongh, A., Ten Broeke, E., & Renssen, M. R. (1999). Treatment of specific phobias with eye movement desensitization and reprocessing (EMDR): Protocol, empirical status, and conceptual issues. *Journal of Anxiety Disorders, 13,* 69–85.

Derogatis, L. R. (1992). *The Brief Symptom Inventory: Administration, scoring and procedures manual—II* (2nd ed.). Baltimore: Clinical Psychometric Research.

DeRubeis, R., & Crits-Christoph, P. (1988). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting and Clinical Psychology, 66,* 37–52.

Feske, U. (1998). Eye movement desensitization and reprocessing treatment for posttraumatic stress disorder. *Clinical Psychology: Science and Practice, 5,* 171–181.

Feske, U., & Goldstein, A. J. (1997). Eye movement desensitization and

reprocessing treatment for panic disorder: A controlled outcome and partial dismantling study. *Journal of Consulting and Clinical Psychology, 65,* 1026–1035.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). *Structured Clinical Interview for DSM–IV Axis I disorders, patient edition.* New York: Biometrics Research Department, New York State Psychiatric Institute.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., & Benjamin, L. (1994). *Structured Clinical Interview for DSM–IV Axis II personality disorders.* New York: Biometrics Research Department, New York State Psychiatric Institute.

Foley, T., & Spates, C. R. (1995). Eye movement desensitization of public-speaking anxiety: A partial dismantling. *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 321–329.

Gelder, M. G., Bancroft, J. H. J., Gath, D. H., Johnston, D. W., Mathews, A. M., & Shaw, P. M. (1973). Specific and non-specific factors in behavior therapy. *British Journal of Psychiatry, 123,* 445–462.

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.

Goldstein, A. J., & Feske, U. (1994). Eye movement desensitization and reprocessing for panic disorder: A case series. *Journal of Anxiety Disorders, 8,* 351–362.

Hollingshead, A. B. (1975). *Four factor index of social status.* (Available from A. B. Hollingshead, Department of Sociology, Yale University, P.O. Box 1965, New Haven, CT 06520)

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.

Lipke, H. (1997). Commentary on the Bates et al. report on eye-movement desensitization and reprocessing (EMDR). *Journal of Anxiety Disorders, 11,* 599–602.

Marcus, S., Marquis, P., & Sakai, C. (1997). Controlled study of treatment of PTSD using EMDR in an HMO setting. *Psychotherapy, 34,* 307–315.

Margraf, J., & Schneider, S. (1990). *Panic* (2nd ed.). Berlin: Springer Verlag.

Marks, I. M., Swinson, R. P., Başoğlu, M., Kuch, K., Noshirvani, H., O'Sullivan, G., Lelliott, P. T., Kirby, M., McNamee, G., Sengun, S., & Wickwire, K. (1993). Alprazolam and exposure alone and combined in panic disorder with agoraphobia: A controlled study in London and Toronto. *British Journal of Psychiatry, 162,* 776–787.

Marquis, J. H. (1991). A report on 78 cases treated by eye movement desensitization. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 319–323.

Mavissakalian, M. (1987). The placebo effect in agoraphobia. *Journal of Nervous and Mental Disease, 175,* 95–99.

McGlynn, F. D. (1997). Response to Lipke's comment. *Journal of Anxiety Disorders, 11,* 603–606.

Muris, P., & Merckelbach, H. (1997). Treating spider phobics with eye movement desensitization and reprocessing: A controlled study. *Behavioural and Cognitive Psychotherapy, 25,* 39–50.

Muris, P., Merckelbach, H., Holdrinet, I., & Sijsenaar, M. (1998). Treating phobic children: Effects of EMDR versus exposure. *Journal of Consulting and Clinical Psychology, 66,* 193–198.

Muris, P., Merckelbach, H., van Haaften, H., & Mayer, B. (1997). Eye movement desensitization and reprocessing versus exposure in vivo: A single-session crossover study of spider-phobic children. *British Journal of Psychiatry, 171,* 82–86.

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

Rogers, J. L., Howard, K. I., & Vessey, J. T. (1993). Using significance tests to evaluate equivalence between two experimental groups. *Psychological Bulletin, 113,* 553–565.

Rothbaum, B. O. (1997). A controlled study of eye movement desensitization and reprocessing in the treatment of posttraumatic stress disordered sexual assault victims. *Bulletin of the Menninger Clinic, 61,* 317–334.

Scheck, M. M., Schaeffer, J. A., & Gillette, C. (1998). Brief psychological intervention with traumatized young women: The efficacy of eye movement desensitization reprocessing. *Journal of Traumatic Stress, 11,* 25–44.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures.* New York: Guilford Press.

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.

Shear, M. K., & Maser, J. D. (1994). Standardized assessment for panic disorder research. *Archives of General Psychiatry, 51,* 346–354.

Steketee, G., & Goldstein, A. J. (1994). Reflection on Shapiro's reflections: Testing EMDR within a theoretical context. *Behavior Therapist, 17,* 156–157.

Taylor, S., Woody, S., Koch, W. J., McLean, P., Paterson, R. J., & Anderson, K. W. (1997). Cognitive restructuring in the treatment of social phobia. *Behavior Modification, 21,* 487–511.

Telch, M., Brouillard, M., Telch, C. F., Agras, W. S., & Taylor, C. B. (1989). Role of cognitive appraisal in panic-related avoidance. *Behaviour Research and Therapy, 27,* 373–383.

van Balkom, A. J. L. M., Bakker, A., Spinhoven, P., Blaauw, B. M. J. W., Smeenk, S., & Ruesink, B. (1997). A meta-analysis of the treatment of panic disorder with or without agoraphobia: A comparison of psychopharmacological, cognitive–behavioral, and combination treatments. *Journal of Nervous and Mental Disease, 185,* 510–516.

van Etten, M., & Taylor, S. (1998). Comparative efficacy of treatments for posttraumatic stress disorder: A meta-analysis. *Clinical Psychology and Psychotherapy, 5,* 126–145.

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry, 33,* 1111–1115.

Williams, K. E., & Chambless, D. L. (1990). The relationship between therapist characteristics and outcome of in vivo exposure treatment for agoraphobia. *Behavior Therapy, 21,* 111–116.

Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR) treatment for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology, 63,* 928–937.

Received July 14, 1999
Revision received February 22, 2000
Accepted March 8, 2000 ■