Journal of Consulting and Clinical Psychology
1995, Vol. 63, No. 6, 928-937

0022-006X/95/$3.00

SPECIAL FEATURE

Eye Movement Desensitization and Reprocessing (EMDR) Treatment for Psychologically Traumatized Individuals

Sandra A. Wilson
The Union Institute

Lee A. Becker
University of Colorado, Colorado Springs

Robert H. Tinker
Colorado Springs, Colorado

The effects of 3 90-min eye movement desensitization and reprocessing (EMDR) treatment sessions on traumatic memories of 80 participants were studied. Participants were randomly assigned to treatment or delayed-treatment conditions and to 1 of 5 licensed therapists trained in EMDR. Participants receiving EMDR showed decreases in presenting complaints and in anxiety and increases in positive cognition. Participants in the delayed-treatment condition showed no improvement on any of these measures across the 30 days before treatment, but after treatment participants in the delayed-treatment condition showed similar effects on all measures. The effects were maintained at 90-day follow-up.

Eye movement desensitization and reprocessing (EMDR) is a controversial treatment that claims to resolve long-standing traumatic memories within a few treatment sessions. During EMDR treatment, the client is asked to hold in mind an image of the trauma, a negative self-cognition, negative emotions, and related physical sensations about the trauma. While doing so, the client is instructed to move her or his eyes quickly and laterally back and forth for about 15 to 20 s, following the therapist's fingers. Other forms of left-right alternating stimulation (auditory, tactile) are sometimes used (Shapiro, 1995). The client then reports the images, cognitions, emotions, and physical sensations that emerged. This recursive procedure continues until desensitization of troubling material is complete and positive self-cognitions have replaced the previous negative self-cognition.

The initial report by Shapiro (1989b) indicated that EMDR markedly reduced anxiety associated with a traumatic memory and increased the perceived validity of positive cognitions within a single session. Shapiro used 22 participants with combat or sexual trauma, who were randomly assigned to one session of EMDR or to a control condition. Control participants were asked to recall the traumatic memory, but they did not perform the eye movements. Shapiro reported that EMDR led to significant and enduring positive behavioral changes, as rated by the participants and their significant others.

Case reports support that EMDR is efficacious in treating traumatic memories (Kleinknecht & Morgan, 1992; Levin, 1993; Lipke & Botkin, 1992; Marquis, 1991; McCann, 1992; Page & Crino, 1993; Pellicer, 1993; Puk, 1991; Shapiro, 1989a; Spector & Huthwaite, 1993; Wernick, 1993; Wolpe & Abrams, 1991). However, a controlled, single-case study has reported negative results (Acierno, Tremont, Last, & Montgomery, 1994). Overall, these case reports provide clinical observations and generate hypotheses for future research but lack the rigor needed for empirical evaluation.

Although Shapiro (1989b) reported positive results, critical reviews have detailed a number of methodological shortcomings: There were no pre- and posttreatment objective or standardized measures used to assess treatment, there was no objective or standardized PTSD diagnosis, the research design did not control for nonspecific treatment (placebo) effects or therapist demand characteristics, the novelty and complexity of the EMDR treatment was not controlled for, and the sample size

Sandra A. Wilson, School of Clinical Psychology, The Union Institute; Lee A. Becker, Department of Psychology, University of Colorado at Colorado Springs; Robert H. Tinker, independent practice, Colorado Springs, Colorado.

Special thanks go to the research team of Michael Galvin, Craig Gillette, Jim Knipe, Laura Knutson, and Beverly Schoninger; consultants Bob Atkins, Sandra Foster, John Hartung, Ethel Kloos-Fenn, and Andrew Sweet; and secretarial assistants Sherry Marshall, Donna Matthews, and Lea Rosenberry. We also thank the participants and their families for their generous participation and support.

Portions of this article were presented in August 1994 at the 102nd Annual Convention of the American Psychological Association, Division 12, and in November 1994 at the 10th annual meeting of the International Society for Traumatic Stress Studies.

Correspondence concerning this article should be addressed to Lee A. Becker, 1420 Austin Bluffs Parkway, P.O. Box 7150, Department of Psychology, University of Colorado, Colorado Springs, Colorado 80933-7150. Electronic mail may be sent via Internet to lbecker@mail.uccs.edu.

was small (Acierno, Hersen, Van Hasselt, Tremont, & Meuser, 1994; Herbert & Meuser, 1992; Lohr et al., 1992).

Other controlled studies of EMDR can be faulted for using small sample sizes (Boudewyns et al., 1993; Jensen, 1994; Pitman et al., 1993), limited or atypical populations (Boudewyns et al., 1993; Jensen, 1994; Pitman et al., 1993; Sanderson & Carpenter, 1992), or treatment integrity issues (Jensen, 1994; Pitman et al., 1993; Sanderson & Carpenter, 1992). Considering the methodological shortcomings of EMDR studies, Acierno et al. stated that ". . . failure to detect significant change following EMDR does not establish its lack of efficacy" (1994, p. 296).

The present study investigated the effects of EMDR on traumatic memories and psychological symptoms. A traumatic memory was defined as a memory of a stressful event that continues to produce PTSD symptoms for the individual. There were four hypotheses: (a) the presenting complaints associated with the traumatic memory, such as nightmares, flashbacks and panic, would be reduced or eliminated as a result of EMDR; (b) that EMDR would reduce anxiety related to the trauma; (c) that EMDR would increase positive self-cognition related to the traumatic event; and (d) that the treatment effect would be maintained at the 90-day follow-up. This study strives to be responsive to the issues raised in critical reviews (Acierno et al., 1994; Herbert & Meuser, 1992; Lohr et al., 1992): a large and diverse sample ($N = 80$) was used; participants were randomly assigned to treatment or control conditions, and to one of five EMDR-trained therapists; pre- and posttreatment evaluations were conducted by an independent assessor, using objective and standardized measures; participants engaged in no other therapy while they were in EMDR treatment; PTSD diagnoses were objectively made; and treatment was monitored.

Method

Participants

There were 40 female and 40 male participants, ranging in age from 21 to 63 years ($M = 39$) who were experiencing traumatic memories. Their education ranged from 10 to 24 years ($M = 15$ years); income ranged from $5,000 to over $50,000 ($Mdn = $20,000); 39% were married, 24% were single, 21% were divorced, 10% were cohabitating, 4% were separated, and 2% were widowed; 96% were White, and 4% were Hispanic.

The major trauma categories were physical–mental abuse (26%), death of a significant other (19%), rape and sexual molestation (22%), relationship crisis (14%), health crisis (9%), phobic memory (6%), and combat trauma (4%). The trauma itself had occurred from 3 months to 54 years ($Mdn = 13.5$ years) before the beginning of the study. Twenty-nine percent of the participants were involved in outpatient therapy just before the onset of the study, another 35% had previously been in therapy, and the remaining 36% had never been in therapy.

Participants were recruited through announcements seeking people with traumatic memories that were placed in local newspapers, victim assistance agencies, local colleges, and therapists' offices. We conducted an initial telephone screening for the 207 people who responded to the announcements to determine whether they met the basic criterion of having a traumatic memory that was interfering with their life (e.g., flashbacks, nightmares, avoidant behaviors, increased anger, or irritability). The 115 people who met the criterion were invited to a more intensive, individually administered, selection interview. The selection interview (administered by Sandra A. Wilson) confirmed that the potential participant met the inclusion criterion and did not meet any of the exclusion criteria. People were excluded from the study if they were experiencing specific medical, psychological, or legal problems. Medical exclusion criteria included vision problems, epilepsy, pregnancy, or neurological impairment. Psychological exclusion criteria included psychosis, dissociative disorders, active substance abuse, or active suicidal ideation. Persons involved in current legal proceedings resulting from the traumatic event or receiving disability benefits as a result of the trauma were also excluded. Eighty people were selected to participate in the study. Six participants dropped out early in the study and were replaced. Three participants dropped out because of medical problems that interfered with scheduling appointments, two dropped out because of scheduling problems that were due to employment, and one participant was dropped from the study after she found that she was pregnant.

Participants in the study were asked to make a commitment not to attend other psychotherapy until after the treatment phase of the study. They were informed that they could return to their regular therapy during the 90-day follow-up period but that they should refrain from dealing with the specific traumatic memory treated in the present study. The exit interviews confirmed that none of the participants received other therapy during the treatment phase of the study. Some participants did return to therapy during the follow-up period, but all indicated compliance with the requirement to refrain from dealing with the traumatic memory treated in this study.

Some participants were assigned a posttraumatic stress disorder diagnosis based on the Post-Traumatic Stress Disorder Interview (PTSD-I; Watson, Juba, Manifold, Kucala, & Anderson, 1991) administered at the initial pretest session. Forty-six percent of the participants ($n = 37$) met all five of the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994) diagnostic criteria for PTSD. Of those participants who were not assigned a PTSD diagnosis ($n = 43$), 30% met criterion A (history of the trauma), 93% met criterion B (memory intrusions), 63% met criterion C (avoidance of the memory); 54% met criterion D (hyperarousal), and 100% met criterion E (the symptoms lasted for at least 30 days). Overall, 75% of all the participants met at least four of the five PTSD criteria; 94% met at least three of the criteria, 95% met at least two of criteria, and 100% met at least one of the criteria.

Therapists and Independent Assessor

Four of the five therapists had received advanced EMDR training; the other therapist had received introductory-level EMDR training. EMDR was administered by two female and three male licensed therapists in private practice whose EMDR experience ranged from several months to 3 years. Their general clinical experience ranged from 5 years to 27 years. Three were licensed psychologists, one was a licensed psychotherapist, and one was a licensed social worker. The independent assessor was a licensed psychologist with little previous knowledge of EMDR and reported being skeptical at the beginning of the study.

Process Measures

Process measures were taken by the therapists. EMDR focuses on the Subjective Units of Disturbance Scale (SUDS) and Validity of Cognition Scale (VOC) ratings as measures of therapy progress. Therefore the SUDS and VOC scores collected by the therapists were considered to be treatment process measures rather than treatment outcome measures.

SUDS. The SUDS, taken from the Subjective Anxiety Scale

(Wolpe, 1990), is a single-item measure of anxiety experienced while thinking about a particular traumatic event. It ranges from 0 (*neutral*) to 10 (*the highest level of disturbance imaginable*).

VOC. The VOC (Shapiro, 1989b) is a single-item measure of the validity of alternative positive cognitions. The therapist first elicits from the participant a negative cognition associated with the traumatic event and then asks what the participant would rather believe about himself or herself. The VOC measures the degree of acceptance of that positive cognition on a scale that ranges from 1 (*completely untrue*) to 7 (*completely true*). The process of choosing the alternative positive cognition often involves considerable interaction between the therapist and participant. For that reason, we decided that the VOC measures should not be collected by the independent assessor.

Outcome Measures

All outcome measures were administered by the independent assessor.

SUDS. SUDS (Wolpe, 1990) measures collected by the independent assessor were treated as outcome measures.

Impact of Event Scale (IES). The IES (Horowitz, Wilner, & Alvarez, 1979) assesses symptoms of avoidance and intrusions experienced over the previous week that were related to a particular stressful event.

State–Trait Anxiety Inventory (STAI). The STAI (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) includes two scales that measure the amount of anxiety that a person "generally feels" (trait anxiety) and the amount felt "right now" (state anxiety).

Symptom Check List (SCL-90-R). The *SCL-90-R* (Derogatis, 1992) measures the occurrence of psychological symptoms for psychiatric and medical patients. Because of their relevance to this study, the following dimensions were analyzed: Somatization, Interpersonal Sensitivity, Depression, and Anxiety.

Measurement Time One (T1): Pretreatment

The pretreatment measurements occurred approximately 1 week after the selection interview. Each participant was individually assessed. As participants had not yet been assigned to the treatment condition or to a therapist, both the independent assessor and the participants had not been informed of those conditions. See Table 1 for a diagram of the experimental design.

At the conclusion of the pretreatment assessment, each participant met with the principal investigator, who explained that because of the large number of people participating in the research, not everyone could begin treatment immediately and that assignment to starting dates was done randomly. Participants were then randomly assigned to treatment conditions and to one of the therapists.

Table 1
The Experimental Design

	Measurement time				
Treatment condition	T1	T2	T3	T4	T5
EMDR treatment	O1	X O2	90 days	O3	
Delayed EMDR treatment	O1	O2 X	O3	90 days	O4

Note. T = time of measurement; EMDR = eye movement desensitization and reprocessing; O = observation; X = treatment administered.

EMDR Treatment Procedure

EMDR, individually administered during three 90-min sessions, included six phases: (a) preparation, (b) baseline assessment, (c) desensitization, (d) installation of the positive cognition, (e) body scan, and (f) closure.

The preparation phase included a description of EMDR and a discussion of the traumatic event that the participant had experienced. In addition, the participant was asked to formulate an image of a place that provided a feeling of calm and safety that could be used, if needed, as a respite during the course of the treatment. During baseline assessment, the trauma became the focus of treatment. We asked the participant to read aloud her or his description of the trauma, to visualize a picture of it, and to think about the negative cognition associated with the trauma. Next, we asked the participant to think about an alternate positive cognition and to rate its validity on the VOC scale. The participant was then asked to remember the trauma including the image, the negative cognition, the emotions, and bodily sensations, and to rate the distress he or she felt using the SUDS scale. During the desensitization phase, after the SUDS and VOC ratings were given, the therapist engaged the participant in the eye movements while the participant held in mind the image, cognition, emotions, and physical sensations. The average set of eye movements was 24–60 traverses, lasting from 12–30 s. At the end of each set the therapist asked the participant to take a deep breath and then asked what emerged. The therapist then helped to guide the participant's focus before beginning the next set of eye movements. This process continued for approximately 60 min, with several checks on the SUDS and VOC scales. When the SUDS score was low (0–3) and the VOC score was high (5–7), the therapist completed an "installation" by requesting the participant to pair the positive cognition with the memory of the trauma and perform an additional set of eye movements. This phase concluded with the therapist taking measures on the SUDS and VOC scales. At closure, if the participant reported being upset at the conclusion of the session, relaxation techniques and eye movements were used until the participant reported feeling calmer. The participant was informed that the processing of the trauma might continue and to keep a written journal describing what emerged.

Measurement Times Two (T2), Three (T3), Four (T4), and Five (T5)

The time interval between T1 and T2 and between T2 and T3 was 1 month. Measures collected at T2 served as the posttreatment assessment for participants in the EMDR condition and as a second pretreatment assessment for participants in the delayed-EMDR condition. Participants in the delayed-EMDR condition received EMDR between T2 and T3; measures collected at T3 served as the posttreatment assessment for those participants. The follow-up occurred 90 days after the posttreatment assessment; at T4 for participants in the EMDR condition and at T5 for participants in the delayed-EMDR condition. The measures and procedures used at T2, T3, T4, and T5 were the same as those used at T1.

Results

Reliability of PTSD Diagnoses

We diagnosed PTSD using the PTSD-I (Watson et al., 1991) requiring clinical judgment to assess the history section (criterion A) of the PTSD criteria. Three clinicians independently judged the criterion A responses of each participant using the *DSM–IV* (American Psychiatric Association, 1994) criteria. The interrater agreements (kappa; Cohen, 1977) for those

judgments were >.84. Disagreements were resolved by discussion among the three raters. The three symptom criteria (criteria B, C, and D) are measured by 17 7-point rating scales that are objectively scored.[1]

Equivalence of Condition

Pretreatment differences between participants assigned to the treatment and delayed-treatment conditions were analyzed by *t* tests (age, duration of the trauma, and the dependent variables) or chi-square tests (marital status, education, income, and current or previous therapy status). There were no significant differences, which indicated that the participants in the two conditions were demographically similar and that there were no preexisting differences on the dependent measures.

Treatment Integrity

Standardized treatment protocols were provided for every session for each participant. The protocol included 11 sequential steps to be followed and recorded by the therapist. The therapists completed 98% of the 2,640 data elements requested in the protocols. All participants met the protocol guidelines for the number of treatment sessions attended. The length of the treatment session was sometimes shortened if the SUDS reached 0 and the VOC reached 7, and the participant and therapist both agreed that the trauma had been resolved. The mean length of time that a participant spent in treatment across all three sessions was 207 min. Therapist drift (Kazdin, 1994) was minimized by a series of research team meetings held before the start of the research in which the treatment protocol was reviewed and by daily supervision of each therapist by the principal investigator.

Process Measures

The therapist measured the progress of the EMDR therapy by assessing the participant's anxiety associated with the traumatic memory (SUDS) and the participant's acceptance of the positive cognition with regard to the trauma (VOC) at the beginning and end of each session. We analyzed the SUDS scores within each of the three sessions using 2 (beginning vs. end of session) × 2 (EMDR vs. delayed EMDR) analyses of variance (ANOVAs). For each of the three sessions, the SUDS scores significantly decreased from the beginning to the end of the session: Session 1, $F(1, 78) = 161.27$; Session 2, $F(1, 78) = 73.43$; Session 3, $F(1, 78) = 68.38$; all $ps < .001$ (refer to Figure 1). In Session 1, the change in SUDS scores from the beginning to the end of the session was greater for the EMDR condition than for the delayed-EMDR condition, $F(1, 78) = 7.85$, $p < .01$. That interaction did not occur in any of the later sessions, and there were no treatment condition main effects in any session. As shown in Figure 1, the SUDS scores at the beginning of Sessions 2 and 3 were higher (Tukey honestly significant difference [HSD], $p < .05$) than at the end of the previous session.

We also analyzed the VOC scores within each of the three sessions using 2 (beginning vs. end of session) × 2 (EMDR vs. delayed EMDR) ANOVAs. For each of the three sessions, the VOC scores significantly increased from the beginning to the end of the session: Session 1, $F(1, 78) = 160.74$; Session 2, $F(1, 78) = 51.82$; Session 3, $F(1, 78) = 38.16$; all $ps < .001$ (see Figure 1). In Session 1, the overall VOC scores were higher for the EMDR condition than for the delayed-EMDR condition, $F(1, 78) = 5.46$, $p < .05$. That effect did not occur in any of the later sessions and there were no interactions. The VOC score at the beginning of Session 3 was lower than the VOC score at the end of the previous session (Tukey HSD, $p < .05$). This effect was not significant between Sessions 1 and 2 (Tukey HSD, $p > .05$).

Outcome Measures

The outcome measures were examined by three sets of analyses. The first set addressed the effectiveness of treatment relative to the control participants in the delayed-treatment condition. The second set addressed the effectiveness of treatment for the delayed-treatment participants. The third set addressed maintenance at the 90-day follow-up. The means, 95% confidence intervals, and normative data for the nine outcome measures across all measurement times are shown in Figure 2. The normative information shown is described later in the Clinical Significance section.

The primary analyses were multivariate analyses of variance (MANOVAs) on the nine dependent measures. The alpha level of .05 was used for all multivariate tests. Alpha levels for a posteriori univariate tests were set to .006 (using the modified Bonferroni procedure) when the main concern was protection against Type I error and to .05 when the main concern was protection against Type II error.

Treatment participants versus delayed-treatment control participants. We analyzed the effectiveness of the EMDR treatment rel-

[1] The PTSD-I (Watson et al., 1991) was designed to measure the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., revised; American Psychiatric Association, 1987) criteria for PTSD. The *DSM-IV* (American Psychiatric Association, 1994) made one change from the *DSM-III-R* with respect to the 17 items that measure criteria B, C, and D. Item D-6 in the *DSM-III-R* was moved to criterion B in the *DSM-IV*. The other change was a clarification of criterion A, the history of the trauma. In this study, the PTSD diagnosis was made according to the *DSM-IV*. criteria.

The items on the PTSD-I are worded in a way that makes them unsuitable for measuring change in symptoms. For example, many of the items ask if the respondent has "ever" experienced a PTSD symptom associated with the trauma. In an attempt to measure changes in symptoms as a result of treatment, we reworded the items to assess the occurrence of symptoms "within the last 7 days." We chose that time frame because it is used on the IES and the *SCL-90-R*. This revision of the PTSD-I was used at all measurement times after T1. The rewording resulted in a significant drop in the percentage of participants who were diagnosed as having PTSD at T2 for participants in the delayed-EMDR condition. Because none of the other outcome measures showed a comparable decrease at T2 for the delayed-EMDR condition, it was concluded that the revised wording invalidated the scale as a PTSD diagnostic instrument. Therefore, we have not presented posttreatment PTSD diagnostic data based on the PTSD-I. The revised PTSD-I scale does seem to be reliable as an outcome measure (Wilson, Tinker, Becker, & Gillette, 1994). The PTSD-I effect size at measurement T2 was 1.28.

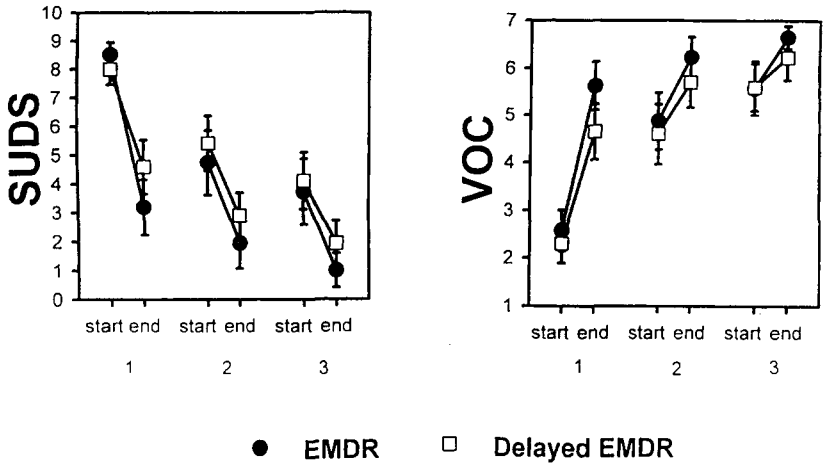

Figure 1. Changes on process measures for treatment and delayed-treatment participants collected by therapist at the start and the end of treatment sessions 1, 2, and 3. Dark symbols represent EMDR treatment group; open symbols represent delayed treatment group. SUDS = Subjective Units of Disturbance Scale; VOC = Validity of Cognition Scale; EMDR = eye movement desensitization and reprocessing. Error bars represent 95% confidence intervals.

ative to the control condition by looking at the change from measurement time T1 to T2 (refer to Table 1). The participants in the treatment condition were treated with three EMDR sessions between those two measurement times, and participants in the delayed condition received no treatment. We performed a multivariate test (Wilks's lambda) using the nine outcome measures as dependent variables. The independent variables were measurement time (T1 vs. T2, a within-subjects factor) and treatment condition (EMDR vs. delayed EMDR, a between-subjects factor). The two-way interaction was significant, Wilks's λ (9, 64) = .40, $p < .0005$. Multivariate simple main effects tests were performed to test specific hypotheses. As expected, participants in the EMDR treatment condition showed improvement, Wilks's λ (9, 64) = .19, $p < .0005$. The canonical correlation for this effect was .90, indicating that 81% of the variability in the discriminant scores for that factor was accounted for by the EMDR treatment. We maintained protection against Type I error for the univariate comparisons by making a Bonferroni adjustment to the alpha level (α = .006). By that criterion, significant improvement was shown on all nine dependent measures. The strength of the relationship between each of the dependent variables and the discriminant function is given by the correlations between them (sometimes called *structure coefficients*). The highest structure coefficients were for the SUDS (−.87) and for the IES Avoidance (−.61) and Intrusion (−.49) scales. The participants in the delayed-EMDR condition showed no change between T1 and T2, Wilks's λ (9, 64) = .86, p = .34. Because no change was expected for the delayed EMDR condition, protection against Type II error was maintained in those univariate comparisons by using an alpha level of .05. Using that criterion, none of the nine measures showed any change. The univariate statistics and structure coefficients for the simple main effect analyses are shown in Table 2.

Treatment effect size. Treatment effect sizes were computed at measurement T2 using Cohen's *d* (Cohen, 1977). Inspection of the effect sizes, shown in Table 2, indicated that each of the effect sizes for the trauma-specific measures (SUDS, IES Intrusion, and IES Avoidance) were larger than those for the general measures (*SCL-90-R* Somatization, Interpersonal Sensitivity, Depression, and Anxiety; State and Trait Anxiety). The composite effect size (Rosenthal & Rubin, 1986) of 1.82 for the trauma-specific measures was greater than the composite effect size of 0.65 for the general measures, $t_c(78) = 7.14$, $p < .001$. Both of the composite effect sizes were significantly different from zero: specific measures, $t_{c(78)=7.49}$, $p < .001$; general measures, $t_{c(78)=2.88}$, $p < .005$.

Effectiveness for delayed-treatment participants. The effectiveness of the EMDR treatment for participants in the delayed-treatment condition was tested by comparing their change from pretreatment (T2) to posttreatment (T3) with the EMDR condition change from pretreatment (T1) to posttreatment (T2; see Table 1). A multivariate test (Wilks's lambda) was performed on the nine outcome measures using those measurement times (pretreatment vs. posttreatment, a within-subjects factor) and treatment condition (EMDR vs. delayed EMDR, a between-subjects factor) as the independent variables. The multivariate test of the pre- to posttreatment difference was significant, Wilks's λ (9, 65) = .16, $p < .0005$. The canonical correlation of .91 indicates that 83% of the variability in the discriminant scores for that factor was accounted for by the EMDR treatment. All the univariate, pre- versus posttreatment main effects were significant (all $ps <$ .006). The multivariate test of the two-way interaction was also significant, Wilks's λ (9, 65) = .76, p = .024. The interaction accounted for 24% of the variability in the discriminant scores. Examination of the univariate tests showed that two measures were significant, the SUDS, $F(1, 73) = 11.86$, $p < .006$, and the IES Avoidance scale, $F(1, 73) = 9.86$, $p < .006$. In both instances, the change was greater for the EMDR condition than for the delayed-EMDR condition. The multivariate test of the treatment condition main effect was not significant, Wilks's λ (9, 65) = .82, p = .13.

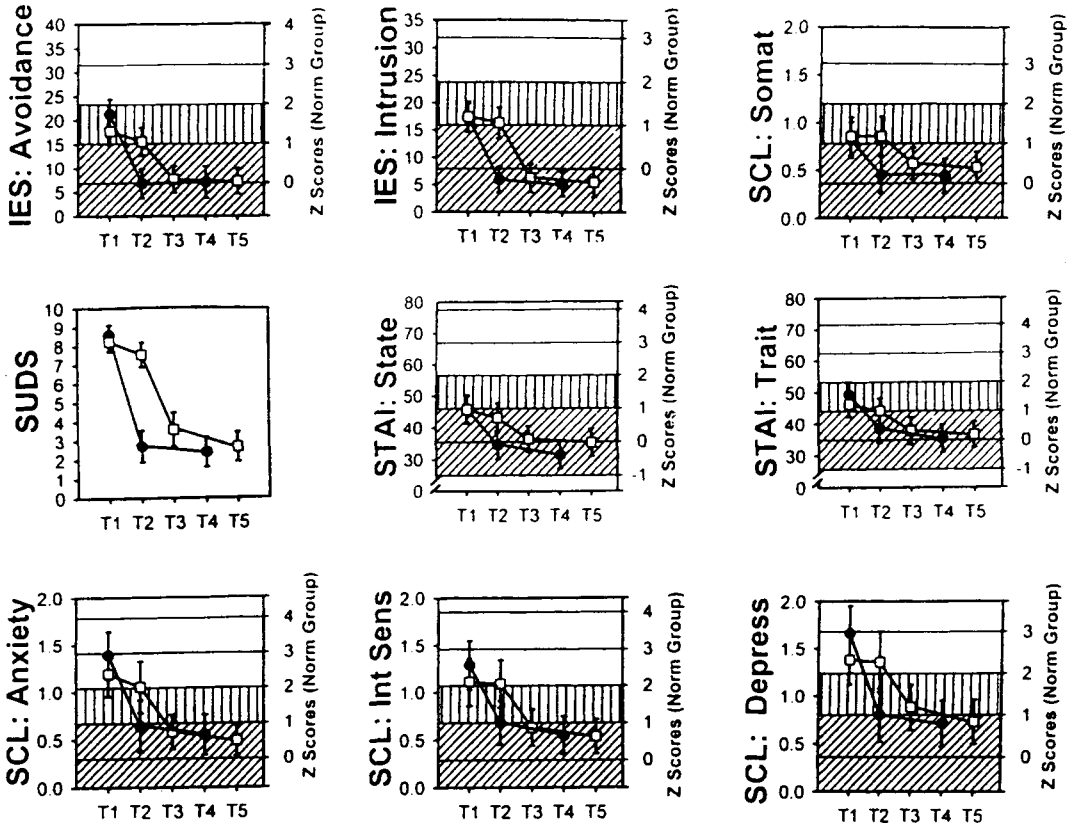

Figure 2. Changes on the outcome measures for treatment and delayed-treatment participants collected by independent assessor. Error bars represent 95% confidence intervals. Treatment occurred for eye movement desensitization and reprocessing (EMDR) participants between T1 and T2. Treatment occurred for delayed-EMDR participants between T2 and T3. Ninety-day follow-up occurred at T4 and T5. The scale on the right vertical axis represents *z* scores based on normative data. The area between ±1.0 *z* scores is filled with diagonal lines, and the area between 1.0 and 2.0 *z* scores is filled with vertical lines. Dark symbols represent EMDR treatment group; open symbols represent delayed treatment group. IES: Avoidance = Impact of Event Scale, Avoidance subscale; IES: Intrusion = Impact of Event Scale, Intrusion subscale; SCL: Somat = *Symptom Check List–90–Revised*, Somatization dimension; SUDS = Subjective Units of Disturbance Scale; STAI: State = State–Trait Anxiety Index, State Anxiety; STAI: Trait = State–Trait Anxiety Index, Trait Anxiety; SCL: Anxiety = *Symptom Check List–90–Revised*, Anxiety dimension; SCL: Int Sens = *Symptom Check List–90–Revised*, Interpersonal Sensitivity dimension; SCL: Depress = *Symptom Check List–90–Revised*, Depression dimension; Norm Group = normative group; T1 through T5 = measurement times 1 through 5.

Ninety-day follow-up. We tested maintenance by comparing the posttreatment scores (T2 for the EMDR condition and T3 for the delayed-EMDR condition) with their 90-day follow-up scores (T4 and T5; see Table 1). We performed a multivariate test on the nine outcome measures using measurement times (posttreatment vs. 90-day follow-up, a within-subjects factor) and treatment condition (EMDR vs. delayed EMDR, a between-subjects factor) as the independent variables. The alpha level for a posteriori univariate tests of the measurement times factor was set at .05 to control for Type II error in order to detect either a worsening of symptoms or a continued improvement in symptoms. There were no significant multivariate effects indicating that the improvement found in the outcome measures at the posttest was maintained at the 90-day follow-up. Although the multivariate test of the posttreatment vs. 90-day follow-up difference was not significant, Wilks's λ (9, 64) = .87, p = .41, the univariate tests showed additional improvement at 90 days for three of the measures: SUDS, $F(1, 72) = 4.46$, p = .04; *SCL–90–R* Interpersonal Sensitivity, $F(1, 72) = 4.85$, p = .03; and Trait Anxiety, $F(1, 72) = 6.11$, p = .02. None of the scales showed a worsening of symptoms at the 90-day follow-up.

Clinical Significance

We examined clinical significance using an approach detailed by Kendall and Grove (1988) where treatment effects are assessed in terms of the mean and standard deviation of a normative comparison group. Normative population data were available for all scales except the SUDS. The normative population

Table 2
Simple Main Effect Fs (T1 vs. T2), Structure Coefficients, and Effect Sizes for the Outcome Measures

Outcome measure	Treatment condition: EMDR	Treatment condition: Delayed EMDR	Cohen's *d*
Presenting complaints			
IES: Avoidance	116.44 (−.61)**	3.12	1.03
IES: Intrusion	74.97 (−.49)**	0.85	1.35
SCL-90-R: Somatization	17.12 (−.23)**	0.14	0.66
Anxiety			
SUDS	235.74 (−.87)**	3.53	2.07
State Anxiety	39.55 (−.36)**	2.07	0.63
Trait Anxiety	54.69 (−.42)**	1.77	0.44
SCL-90-R: Anxiety	46.13 (−.39)**	1.53	0.49
Positive self-evaluation			
SCL-90-R: Interpersonal Sensitivity	15.41 (−.31)**	0.00	0.56
SCL-90-R: Depression	54.73 (−.42)**	0.05	0.62

Note. The delayed EMDR condition received no treatment between T1 and T2. The *df*s for all the univariate *F*s are 1 and 72. Structure coefficients are shown in parentheses. Cohen's *d*s were computed at T2. EMDR = eye movement desensitization and reprocessing; IES: Avoidance = Avoidance subscale of the Impact of Event-Scale; IES: Intrusion = Intrusion subscale of the IES; *SCL-90-R:* Somatization = *Symptom Check List—90—Revised,* Somatization dimension; SUDS = Subjective Units of Disturbance Scale; State Anxiety = State-Trait Anxiety Index, State Anxiety subscale; Trait Anxiety = State-Trait Anxiety Index, Trait Anxiety subscale; *SCL-90-R:* Anxiety = *SCL-90-R* Anxiety dimension; *SCL-90-R:* Interpersonal Sensitivity = *SCL-90-R* Interpersonal Sensitivity dimension; *SCL-90-R:* Depression = *SCL-90-R* Depression dimension.
** $p < .006$.

data for the *SCL-90-R* scales were based on nonpatient normative data ($N = 974$) provided in the *SCL-90-R Administration, Scoring and Procedures Manual* (Derogatis, 1992). The normative population data for the State and Trait Anxiety scales were based on the working adult data ($N = 1,838$) provided in the STAI manual (Spielberger et al., 1983). The IES measures intrusions and avoidances in response to a specific stressful life event, and there are no published "normative comparison group" data. Therefore, for the purposes of the present study, normative population data were based on studies presented in a review of stress response outcome studies by Horowitz, Field, and Classen (1993). The specific studies were selected as normative because Horowitz et al. considered them to be representative of recovery after an acute stress response (20 female survivors of a tornado 68 weeks after the event) or after a successful clinical intervention (35 stress clinic patients 66 weeks after the event) or representative of people who had little stress response at the time of the event or several months later (19 male survivors of a tornado 68 weeks after the event, 37 nonpatient controls for the stress clinic patients 66 weeks after the event, and 15 plane crash rescue workers 82 weeks after the event). The unweighted average of the means and standard deviations for those groups were used as estimates of the normative population means and standard deviations.

The normative comparison group *z* scores are shown on the right vertical axis of each outcome measure in Figure 2. Before treatment, the means of all measures fell near or beyond the normative *z* score of 1.0 (the 84th percentile of the normative comparison group). Three of the *SCL-90-R* scales, Anxiety, Depression, and Interpersonal Sensitivity, were at or beyond the normative *z* score of 2.0 (the 97th percentile of the normative comparison group). These data indicate that the participants in this study experienced symptoms that were more severe than those experienced by most people in the normative group. After treatment, the means went from "deviant" to a within-normal-limits range. At the 90-day follow-up, all means fell within a normative *z* score of ±1.0.

Supplementary Analyses

We performed supplementary analyses to determine whether the improvement from the pretreatment scores to the posttreatment scores could be accounted for by (a) demographic variables, (b) the nature of the trauma (e.g., PTSD diagnosis and type of traumatic event), or (c) the severity of the symptoms as measured by the initial pretest scores. Three multiple linear regression analyses were performed using a composite gain score as the dependent variable. The composite gain score was the discriminant function score that was used to discriminate the pretreatment–posttreatment differences as described in the "Effectiveness for Delayed-Treatment Participants" section of this article. A stepwise procedure was used to solve the multiple linear regressions. Categorical variables were transformed into dummy variables before being included in the regression analysis. The alpha level used for the multiple regression analyses was .05.

We analyzed the following demographic variables: years of education, income, marital status, gender, gender of therapist,

and matching of participant gender and therapist gender. Marital status dummy variables were created for categories that contained 10% or more of the participants. The marital status variables were single (24%), married (39%), and divorced (21%). The remaining marital status categories became the reference category. The gender-matching variable was included because, as Beutler, Machado, and Neufeldt (1994) pointed out, there is a lack of clear evidence on the effects of gender matching. Only one variable, married (or not), predicted the gain in the multiple regression equation (multiple $R^2 = .05, p = .046$). The gain for married people was higher than for those who were unmarried ($\beta = -.23$).

The trauma was characterized by whether a PTSD diagnosis was made for the participant, the nature of the specific trauma event, the duration of the traumatic memory, and therapy status. Dummy variables were created for traumatic events categories that contained 10% or more of the participants. Those events included physical or mental abuse (26%), sexual molestation and rape (21%), death of a significant other (19%), and problems in a relationship with another person (14%). The remaining categories became the reference category. The therapy status dummy variables included those who had been in therapy previously (35%) and those who were in therapy just before the onset of the study (29%). Participants who had never been in therapy served as the reference category. The multiple regression analysis indicated that none of these variables predicted the gain score.

The severity of the symptoms at the beginning of the study was measured by the initial prettreatment scores (at measurement time T1) on the nine outcome measures. The regression analysis also included the severity of the PTSD symptoms as measured by the sum of the 17 PTSD-I scales measuring PTSD criteria B, C, and D. The multiple regression analysis indicated that none of these symptom severity measures was related to the gain score.

Discussion

The present results suggest that EMDR was effective in decreasing symptoms and anxiety associated with traumatic memory and in increasing positive cognition. EMDR effectiveness was demonstrated on different outcome measures after three 90-min treatment sessions, with the effects being maintained at 90 days after treatment.

The SUDS and VOC ratings improved within each of the three treatment sessions. These findings are consistent with those of Shapiro (1989b) who found significant improvement on these measures in a single treatment session. The present results also buttress Shapiro's findings of significant reductions in presenting complaints and anxiety.

The supplementary analyses suggest that EMDR was comparably effective for a range of traumas and individuals. Apparently the treatment was equally effective whether the trauma was related to sexual assault or molestation, physical or mental abuse, relationship trauma, or death of a significant other. Likewise EMDR worked equally well whether the participant had previous therapy or not and for those diagnosed as PTSD versus those who did not receive that diagnosis. Outcomes also did not vary for longstanding traumas or more recent ones, for severe or less severe traumas, or for gender of therapist or participant.

A number of factors limit conclusions that can be drawn from the present study. Behavioral measures of outcome were not used. Other than PTSD, no diagnoses were made, preventing investigation into the effectiveness of the treatment with other diagnoses or the effects of comorbidity. Individuals with strong secondary gain issues were screened from the present study. Treatment integrity was not evaluated by independent observers. Nonspecific treatment effects may be represented in the present results to an unknown extent. Treatment effectiveness with multiple traumatic memories was not investigated. Therapists were more rigorously monitored than what would be expected in usual clinical practice. The present sample was not representative of minority populations.

The effectiveness of various treatments for PTSD has been reviewed by Blake, Abueg, Woodward, and Keane (1993). Although research is limited, they concluded that exposure therapies have somewhat better research support than the psychodynamic approaches and documented that the typical length of treatment for exposure therapies is 10 to 14 sessions. Although EMDR therapy contains a number of elements that are not typical of exposure therapies, to some degree imaginal exposure is involved in EMDR as the participant is exposed to images, cognition, emotions, and physical sensations of the traumatic event. However, length of treatment in the present research was considerably shorter than what was reported by Blake et al. for other therapies.

Despite the brevity of treatment, the mean effect sizes in the present study range from 0.44 to 2.07, with composite effect sizes of 1.82 for trauma-specific measures and 0.65 for general measures of functioning. These effect sizes compare favorably with well-controlled studies on psychotherapy as reported by Lipsey and Wilson (1993). The clinical significance analysis, using normative comparisons, indicated that EMDR brought the participants to within a normal range on the outcome measures.

Treatment effects were found to be stronger for measures that were specifically related to the trauma than for the more general measures. Trauma-specific measures (SUDS; IES Intrusion and Avoidance scales) contain questions about how the specific traumatic memory disrupts everyday functioning and how the participant avoids reminders of the specific trauma. The more general measures (State and Trait Anxiety; *SCL–90–R* scales) contain items that relate to general psychological functioning rather than to symptoms related to a specific traumatic event. It may be that the EMDR treatment was not of sufficient duration to effect larger changes on the more general measures. Furthermore, some participants may have had more than one traumatic memory. The more general measures of functioning might be resistant to change to the extent that additional traumatic memories were not treated. Blake et al. (1993) have noted that previous studies in PTSD typically only note improvement in the more obvious PTSD symptoms, such as intrusions and avoidance, and that evidence of improvement on more general measures of psychological functioning has been lacking. The finding that EMDR treatment had any effect on the more generalized measures raises the possibility that, as the

traumatic memory was desensitized, the general functioning of the participant improved, with the participant becoming less anxious and depressed, experiencing fewer somatic complaints, and improving in self-esteem.

Demand characteristics were minimized in the present study by employing an independent assessor to collect pre- and post-treatment measurements. Therapist-collected SUDS ratings have been criticized as being overly sensitive to therapist demand (Acierno et al., 1994; Herbert & Meuser, 1992; Lohr et al., 1992). Because the termination of the assessment session was not contingent on the SUDS ratings taken by the independent assessor, and because the participant's personal relationship with the assessor was minimal, it is likely that demand characteristics associated with those SUDS ratings were lessened. Even with those controls for demand characteristics in place, it might be argued that nonspecific, placebo treatment effects played a role in the present outcomes. However, PTSD has been noted to be resistant to placebo effects (Solomon, Gerrity, & Muff, 1992). In addition, when Lipsey and Wilson (1993) compared meta-analytic studies using placebo controls with those without placebo controls, they found that the average placebo effect size was .19. This suggests that, although placebo effects could account for some part of the effect sizes in the present study, it is not likely that they could fully account for the present results, where the composite effect size for the trauma-specific measures was over nine times greater. The present results suggest that EMDR shows promise in the treatment of traumatic memories, although the reasons for its effectiveness are yet to be understood.

References

Acierno, R., Hersen, M., Van Hasselt, V. B., Tremont, G., & Meuser, K. T. (1994). Review of validation and dissemination of eye-movement desensitization and reprocessing: A scientific and ethical dilemma. *Clinical Psychology Review, 14,* 287–299.

Acierno, R., Tremont, G., Last, C. G., & Montgomery, D. (1994). Tripartite assessment of the efficacy of eye-movement desensitization in a multi-phobic patient. *Journal of Anxiety Disorders, 8,* 259–276.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington DC: Author.

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

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

Blake, D. D., Abueg, F. R., Woodward, S. H., & Keane, T. M. (1993). Treatment efficacy in posttraumatic stress disorder. In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 195–226). New York: Plenum Press.

Boudewyns, P. A., Stwertka, S. A., Hyer, L. E., Albrecht, J. W., & Speer, E. V. (1993). Eye movement desensitization and reprocessing: A pilot study. *Behavior Therapist, 16,* 30–33.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences.* San Diego, CA: Academic Press.

Derogatis, L. R. (1992). *SCL-90*: Administration, scoring and procedures manual—II. Baltimore: Clinical Psychometric Research.

Herbert, J. D., & Meuser, K. T. (1992). Eye movement desensitization: A critique of the evidence. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 169–174.

Horowitz, M. J., Field, N. P., & Classen, C. C. (1993). Stress response syndromes and their treatment. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects* (pp. 757–773). New York: Free Press.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of event scale: A measure of subjective distress. *Psychosomatic Medicine, 41,* 209–218.

Jensen, J. A. (1994). An investigation of eye movement desensitization and reprocessing (EMD/R) as a treatment for posttraumatic stress disorder (PTSD) symptoms of Vietnam combat veterans. *Behavior Therapy, 25,* 311–325.

Kendall, P. C., & Grove, W. M. (1988). Normative comparisons in therapy outcome. *Behavioral Assessment, 10,* 147–158.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (pp. 19–71). New York: Wiley.

Kleinknecht, R. A., & Morgan, M. P. (1992). Treatment of post-traumatic stress disorder with eye movement desensitization. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 43–50.

Levin, C. (July–August, 1993). The enigma of EMDR. *Family Therapy Networker*, pp. 75–83.

Lipke, H., & Botkin, A. (1992). Brief case studies of eye movement desensitization and reprocessing with chronic post-traumatic stress disorder. *Psychotherapy, 29,* 591–595.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48,* 1181–1209.

Lohr, J. M., Kleinknecht, R. A., Conley, A. T., dal Cerro, S., Schmidt, J., & Sonntag, M. E. (1992). A methodological critique of the current status of eye movement desensitization (EMD). *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 159–167.

Marquis, J. (1991). A report on seventy-eight cases treated by eye movement desensitization. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 187–192.

McCann, D. L. (1992). Post-traumatic stress disorder due to devastating burns overcome by a single session of eye movement desensitization. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 319–323.

Page, A. C., & Crino, R. D. (1993). Eye-movement desensitization: A simple treatment for post-traumatic stress disorder? *Australian and New Zealand Journal of Psychiatry, 27,* 288–293.

Pellicer, X. (1993). Eye movement desensitization treatment of a child's nightmares: A case report. *Journal of Behavior Therapy and Experimental Psychiatry, 24,* 73–75.

Pitman, R. K., Orr, S. P., Altman, B., Longpre, R. E., Poire, R. E., & Lasko, N. B. (1993, May). *A controlled study of EMDR treatment for post-traumatic stress disorder.* Paper presented at the 146th annual meeting of the American Psychiatric Association, Washington, DC.

Puk, G. (1991). Treating traumatic memories: A case report on the eye movement desensitization procedure. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 149–151.

Rosenthal, T., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin, 99,* 400–406.

Sanderson, A., & Carpenter, R. (1992). Eye movement desensitization versus image confrontation: A single-session crossover study of 58 phobic subjects. *Journal of Behavior Therapy and Experimental Psychiatry, 23,* 269–275.

Shapiro, F. (1989a). Efficacy of the eye movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress Studies, 2,* 199–223.

Shapiro, F. (1989b). Eye movement desensitization: A new treatment

for post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 20,* 211–217.

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

Solomon, S. D., Gerrity, E. T., & Muff, A. M. (1992). Efficacy of treatments for posttramatic stress disorder: An empirical review. *Journal of the American Medical Association, 268,* 633–638.

Spector, J., & Huthwaite, M. (1993). Eye-movement desensitization to overcome post-traumatic stress disorder. *British Journal of Psychiatry, 163,* 106–108.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. D., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory.* Palo Alto, CA: Consulting Psychologists Press.

Watson, C. G., Juba, M. P., Manifold, V., Kucala, T., & Anderson, P. (1991). The PTSD interview: Rationale, description, reliability and concurrent validity of a *DSM–II*–Based technique. *Journal of Clinical Psychology, 47,* 179–188.

Wernick, U. (1993). The role of the traumatic component in the etiology of sexual dysfunctions and its treatment with eye movement desensitization procedure. *Journal of Sex Education and Therapy, 19,* 212–222.

Wilson, S. A., Tinker, R. H., Becker, L. A., & Gillette, C. S. (1994, November). *Using the PTSD-I as an outcome measure.* Paper presented at the 10th annual meeting of The International Society for Traumatic Stress Studies, Chicago, IL.

Wolpe, J. (1990). *The practice of behavior therapy* (4th ed.). New York: Pergamon Press.

Wolpe, J., & Abrams, J. (1991). Post-traumatic stress disorder overcome by eye movement desensitization: A case report. *Journal of Behavior Therapy and Experimental Psychiatry, 22,* 39–43.

Received May 16, 1994
Revision received March 27, 1995
Accepted May 12, 1995 ■