

Eye Movement Desensitization and Reprocessing (EMDR): A Meta-Analysis

Paul R. Davidson and Kevin C. H. Parker
Queen's University

Eye movement desensitization and reprocessing (EMDR), a controversial treatment suggested for posttraumatic stress disorder (PTSD) and other conditions, was evaluated in a meta-analysis of 34 studies that examined EMDR with a variety of populations and measures. Process and outcome measures were examined separately, and EMDR showed an effect on both when compared with no treatment and with therapies not using exposure to anxiety-provoking stimuli and in pre-post EMDR comparisons. However, no significant effect was found when EMDR was compared with other exposure techniques. No incremental effect of eye movements was noted when EMDR was compared with the same procedure without them. R. J. DeRubeis and P. Crits-Christoph (1998) noted that EMDR is a potentially effective treatment for noncombat PTSD, but studies that examined such patient groups did not give clear support to this. In sum, EMDR appears to be no more effective than other exposure techniques, and evidence suggests that the eye movements integral to the treatment, and to its name, are unnecessary.

Eye movement desensitization and reprocessing (EMDR), a novel, popular, and controversial treatment, was introduced by Shapiro (1989a, 1989b) as a new treatment for traumatic memories. EMDR has been advanced as a treatment for a diversity of anxiety disorders, including posttraumatic stress disorder (PTSD; e.g., Rothbaum, 1997), traumatic memories (e.g., Shapiro, 1989a, S. A. Wilson, Becker, & Tinker, 1995), panic disorder (e.g., Feske & Goldstein, 1997), claustrophobia (Lohr, Tolin, & Kleinknecht, 1996), blood and injection phobias (Kleinknecht, 1993), and spider phobia (e.g., Muris & Merckelbach, 1997). A complete procedural description of EMDR is given in Shapiro's (1995) treatment manual.

To evaluate EMDR, questions of both overall effectiveness and of mechanism (whether the eye movements in EMDR are necessary) need to be posed. EMDR's mechanism has been a source of controversy. Although the eye movements are integral to the basic procedure (Shapiro, 1995), some researchers have argued that they are not necessary and that EMDR is best understood as an exposure technique (Foley & Spates, 1995; Lohr, Kleinknecht, Tolin, & Barrett, 1995; Lohr, Tolin, & Lilienfeld, 1998; Sanderson & Carpenter, 1992) with a possible thought-stopping component (Cusack & Spates, 1999). Cahill, Carrigan, and Freuh (1999) concluded that although eye movements may have some effect on the measures taken during sessions as part of the treatment process (referred to in the present article as *process measures*), there is little evidence that eye movements have any impact on outcome measures.

These process measures (ratings of subjective units of distress [SUD] elicited by the fear-provoking images and validity of cognition [VoC] ratings of the new cognitions to be associated with the images) are used to determine when phases of treatment end; see Pitman et al., 1996). EMDR treatment generally continues until both SUD and VoC reach a criterion value (Shapiro, 1995). Because they are integral to the treatment, SUD and VoC are readily available measures and were among the first data reported on EMDR. In some early studies (e.g., Sanderson & Carpenter, 1992; Shapiro, 1989a) the only way to compare groups was to use these two process measures. Some early critics (Acierno, Hersen, Van Hasselt, Tremont, & Mueser, 1994; Herbert & Mueser, 1992) advocated standardized measures of treatment outcome, and most recent studies have included a variety of these (e.g., Feske & Goldstein, 1997; S. A. Wilson et al., 1995). For the present analysis the specific link of SUD and VoC to EMDR is acknowledged, and these process measures are examined separately from outcome measures such as psychometric or physiological assessments.

Several reviews have been unfavorable toward the EMDR literature, criticizing both the measures used to assess effectiveness and the hypotheses raised about the mechanism of EMDR. Some early reviews (Herbert & Mueser, 1992; Lohr et al., 1992) focused on case studies and Shapiro's (1989a) early experiment. The latter was a seminal work but was seen as lacking in standardized outcome measures. Lohr et al. (1992) raised the criticism that progress through the stages of EMDR is conditional on changes in reported SUD and VoC measures, whereas this is not the case in most therapy procedures used as control conditions. They saw this contingency as differentially affecting the process measures in EMDR compared to other conditions. As to EMDR's mechanism, Acierno et al. (1994) concluded on the basis of their review that comparisons of EMDR to the same procedure without eye movements yielded equivalent results. They suggested that the eye movements are unnecessary and that EMDR may be viewed as an imaginal exposure technique. Later reviews done by Lohr and his

Paul R. Davidson and Kevin C. H. Parker, Department of Psychiatry, Queen's University, Kingston, Ontario, Canada.

We acknowledge the ongoing support of this and other research by the Hotel Dieu Hospital, Kingston, Ontario, Canada.

Correspondence concerning this article should be addressed to Paul R. Davidson, Department of Psychiatry, Queen's University, 72 Barrie Street, Kingston, Ontario K7L 3J7, Canada. Electronic mail may be sent to prd@post.queensu.ca.

associates (Lohr et al., 1995; Lohr et al., 1998) also found little evidence that eye movements, or even other methods of lateral stimulation, are necessary.

Other reviews have been far less critical of the procedure. DeBell and Jones (1997) and Feske (1998) argued that, although it needs more research, EMDR is probably an effective treatment for at least some populations. In her review Shapiro (1996a) tended to overlook the weaknesses others have noted and argued for recognition of EMDR as an empirically validated treatment for PTSD. In addition, she has vigorously criticized the critics of the procedure (e.g., Shapiro, 1996b, 1999).

This body of reviews gives rise to four questions that are of interest to the present analysis. First, is EMDR an effective treatment? Chambless and Hollon (1998) outlined criteria for empirically supported therapies. On the basis of these criteria, DeRubeis and Crits-Christoph (1998) described EMDR as a "potentially effective" treatment for noncombat PTSD, although McNally (1999) suggested that Chambless and Hollon's criteria are loose and merit reconsideration. There are plentiful data on EMDR. It has been used in a number of populations and evaluated with process measures in addition to a variety of outcome measures. It has been compared with several types of controls, including waiting list, nonspecific treatment, and exposure (imaginal and *in vivo*). So, the question of effectiveness has several parts. When assessed by meta-analysis, is EMDR an effective treatment? If so, by what sorts of measures, in what populations, and compared to what controls? How large is the effect size?

Second, are the eye movements necessary? Originally described as "the crucial component of the . . . procedure" (Shapiro, 1989a, p. 220), therapist finger movements are more recently seen as part of a group of external alternating stimuli to which a client's attention is directed (Shapiro, 1996b, p. 209). Some studies have shown that alternatives to eye movements, such as bilateral finger tapping, are equally as effective as eye movements. Rather than being seen as disconfirming EMDR, the alternatives used in these studies have been incorporated by Shapiro as valid alternative techniques of EMDR (Shapiro, 1994). Although this may reflect a willingness to adapt the technique to empirical findings, this flexibility has made EMDR a moving target for researchers trying to examine its techniques. This has led DeBell and Jones (1997) and Rosen (1999) to speculate that future researchers will be challenged to find alternatives for comparative research that will not be criticized as a valid form of EMDR. Nevertheless, the question remains: Are eye movements or alternating stimuli necessary?

Third, does it matter who trains the therapists? Some studies with nonconfirmatory results have been dismissed by EMDR proponents who argue that the studies are not a valid test of EMDR if the therapists are not trained by the EMDR Institute (e.g., Greenwald, 1996; Lipke, 1997; Shapiro, 1995, p. 333). Issues of treatment fidelity are important, and we need to examine the data to see if the lack of this training has an effect on EMDR process or outcomes.

Fourth, is EMDR more effective with some disorders than with others? DeRubeis and Crits-Christoph (1998) argued that it is. In suggesting that EMDR is a potentially effective treatment for noncombat PTSD they cited two studies: those of S. A. Wilson et al. (1995) and Renfrey and Spates (1994). They acknowledged problems with their conclusion, noting first that Renfrey and Spates's study yielded null results and does not support their

contention. Second, they noted that, despite the powerful effect of EMDR compared to no treatment demonstrated by S. A. Wilson et al. (1995), the study does not speak clearly about patients with PTSD: Only 42% of the participants met the criteria for PTSD, and their data were not examined separately from the data of the rest of the sample (a group of patients with traumatic memories who did not meet the criteria for PTSD).

In the present article we report a meta-analysis designed to answer the four questions. A meta-analysis is a quantitative examination of the grouped outcomes of several studies (Rosenthal, 1991). Access to aggregate data can allow a different kind of objectivity than does a traditional review article. The EMDR literature seems ripe for such an analysis, but its diversity of populations and measures presents some difficulties. Rather than focusing on a single group or measure in the study of EMDR, in the present analysis we selected all the studies in the literature, with the exception of those that have fatal methodological flaws. We examined the resulting data in many ways to address the questions raised in the reviews cited above. In addition, we included as much raw data as possible in this article so that interested readers can examine the data and our conclusions together.

In this article we aim to quantify the size of the effect of EMDR, combining the results of the many studies in the literature. These studies have compared EMDR with various control and contrast conditions, ranging from no treatment (e.g., S. A. Wilson et al., 1995) to variants of EMDR with one aspect changed (e.g., Cusack & Spates, 1999). For the analysis, we grouped studies by the control or contrast conditions used, and we made an estimate of the size of the effect of EMDR when compared with each type of control condition separately. One group of studies we examined compared EMDR with a variant with no eye movements. Thus, the meta-analysis addresses questions similar to those dealt with by Cahill et al. (1999)—"compared to what is EMDR effective?" and "are the eye movements necessary?"—and adds the questions "what patients will benefit most from EMDR?" and "does it matter who trains the therapists?"

Method

Procedure

This analysis includes only published studies that permitted a clear unconfounded examination of the effect of EMDR and excludes articles only given at conferences. We conducted searches of MedLine and PsycINFO from 1988 to April 2000 and Current Contents from 1997 to March 2000 using *EMDR* and *eye movement desensitization and reprocessing* as text words. All articles produced by the search were considered. We also checked the reference lists in each of the articles for additional relevant studies. Only treatment experiments were included. Because most of the studies did not report follow-up data we chose to restrict our examination to data collected immediately after treatment.

Thirty-four studies were selected for inclusion. One of these studies (Grainger, Levin, Allen-Byrd, Doctor, & Lee, 1997) has been omitted by some other reviewers (e.g., Cahill et al., 1999) because the assignment of participants to conditions, although unsystematic, was not random. However, Grainger et al. argued that there was no systematic bias to group assignment of participants and that the article is unique in that it reports EMDR treatment of survivors of a recent disaster. We elected to include it. Apart from this single exception, all articles that reported group comparisons used randomized assignment. A systematic examination of the quality of the individual studies in terms of blindness of raters, treatment

fidelity, and so on, is beyond the scope of this article. For that sort of examination the reader is referred to the series of reviews conducted by Lohr and his colleagues (Lohr et al., 1995; Lohr, Lilienfeld, Tolin, & Herbert, 1999; Lohr et al., 1998), to the recent reviews done by Cahill et al. (1999) and DeBell and Jones (1997), and to the many works of Shapiro (e.g. Shapiro, 1995, 1996a, 1996b, 1999).

The major focus of this article was the comparison of participants treated with EMDR with other participants treated another way or left untreated (between-subjects designs). Because they have often been reported, a secondary focus was the within-subject effects (pre- to posttest comparisons). Two studies without control groups were included only for the within-subject comparison (Forbes, Creamer, & Rycroft, 1994; Vaughan, Wiese, Gold, & Tarrier, 1994) because they reported clear pre-post comparisons in a clinical sample on a variety of outcome measures for participants treated with EMDR.

Not all published studies were among the 34 included in the analysis. Two exclusions were due to nonrandom assignment of participants (Henry, 1996; Silver, Brooks, & Obenchain, 1995), and one was excluded because treatment condition was confounded with therapist (Rogers et al., 1999). One study (Boudewyns & Hyer, 1996) was excluded because the data reported were insufficient for calculation of effect sizes in which comparisons were not significant and because the authors did not provide additional data when they were requested.

The studies included in the present analysis and their median effect sizes are presented in the Appendix. Many articles reported more than one comparison type; that is, they compared EMDR with more than one control or comparison condition (e.g., waiting list and imaginal exposure). In addition, many articles reported more than one measure of outcome or process. For each article, and within each article for each type of comparison, we examined each measure and, where possible, calculated an effect size estimate. For studies in which more than one effect size was available for a particular comparison we used the median of the values available so that all studies in a particular comparison were equally weighted.

Calculation of Effect Sizes

We estimated all effect sizes using Rosenthal's (1991) formulae for r . The values of r range from -1 to 1 . The absolute value of r is the size of the effect, and the sign indicates the direction, with a positive value indicating that EMDR was more effective than the control and a negative value indicating the control condition was more effective than EMDR. Where summary statistics for the study were not available in the text or tables, we used other approaches to calculations. Where means, standard deviations, and sample sizes were available, we computed between-groups t tests and used them to generate r s. Sometimes, where analysis of variance (ANOVA) statistics were reported, we recomputed statistics so that an F with the first degree of freedom equal to 1 was available to compute r . Where the necessary statistics were unavailable, we asked authors to provide summary statistics sufficient to allow the computation of r s for the entire design.

Two major classes of statistic were computed. Within-subject designs were typically reported using repeated measures ANOVAs or paired difference t tests. Between-subjects designs were typically reported using t tests or ANOVAs. Some articles reported between-subjects results using the main effect between groups in an analysis-of-covariance design with posttest scores as a dependent variable and pretest scores as a covariate.

Measure Type Categories

For the purposes of data organization we assigned the measures used in the various studies to one of six categories, five of which were outcome measures; the other was process measures. These categories were SUD/VoC (process measures), psychometric measures of PTSD (this formed a separate category because EMDR was originally suggested as a treatment

for PTSD [Shapiro, 1989a]), other psychometric measures, physiological measures, behavioral measures (including behavioral avoidance), and other measures. "Other measures" was a category of exclusion and consisted of measures of pain reaction, ratings of vividness of image, and so on. For each article we calculated an effect size for one or more of these six measure categories on the basis of the measures reported in the article. If only one measure in a particular category was used in a study, the effect size for it was included. If more than one measure in a particular measure category was used in a study, the median of the calculated effect sizes for that measure category was used. If there was no measure in a category, no value was recorded for that category. This procedure was followed to avoid giving undue weight to studies with more than one measure in a given category.

Comparison Type Categories

In the 34 studies EMDR was compared with a range of different control conditions. For the purposes of this report we divided these into seven categories. One category was pre-post EMDR, in which pre- and post-EMDR scores were compared within subjects. The other six categories were all comparisons between EMDR and a control or contrast group: (a) no treatment (either waiting list or standard clinic treatment), (b) *in vivo* exposure or cognitive behavior therapy, (c) exposure—not *in vivo* (i.e., in imagination or by video), (d) eyes fixed desensitization reprocessing (EFixDR, in which EMDR procedures were used, omitting only eye movements), (e) other dismantling designs (OthDism, in which EMDR procedures were used except that a single aspect other than eye movements was changed), and (f) nonspecific treatments (these included applied relaxation, biofeedback, active listening, and rapid induction, but not any therapy in which anxiety-provoking stimuli were confronted either imaginably, by video, or *in vivo*). Not all of these conditions were found in any single study, but for each study in which one or more of these conditions was used as a control we calculated an effect size for each comparison condition type separately.

Statistical Analysis

We transformed effect sizes into a normal distribution using Fisher's Z transformation (Rosenthal, 1991, p. 21). We adjusted the total distribution of transformed effect sizes for outlying values and discarded effect sizes that were more than twice the interquartile range from the mean. From a total of 116 effect sizes, 2 (1.7%) outliers were discarded from some analyses but were included in others for the sake of completeness.¹ Although all calculations were done using transformed data, the results (both means and standard deviations) were transformed back and are reported as r s rather than as Z values.

We also report power analyses yielding sample sizes required to achieve an alpha less than or equal to .05 (two tailed) and power greater than or equal to 80%. On the basis of a computed effect size and the underlying variance among the studies it comprises we estimated the sample size required to produce a significant result with sufficient power to avoid

¹ The two data points come from one study: D. L. Wilson et al. (1996). In addition to being statistical outliers, its findings are quite anomalous in the literature. All other studies comparing EMDR with EFixDR show *some* change from pre to post for both treatments. By contrast, in Wilson et al.'s study neither the EFixDR group nor the OthDism group showed any effect of these interventions. This does not seem to be due to an error of selecting untreatable participants, because both control groups were subsequently treated with EMDR and showed an effect almost equal to that of the original group. Although the study strongly supported the effect of EMDR on process measures (no data were reported on outcome measures), it was a questionable study with unusual data, and it was a statistical outlier.

unwarranted retention of the null hypothesis. We report both the number of participants required to achieve this power in a single study if the observed effect size is accurate and the number of studies required to achieve this power in a meta-analysis of studies with the observed effect size and variance. We performed power analysis using GPOWER (Faul & Erdfelder, 1992).

We computed the median effect sizes for each type of comparison in each study and analyzed them separately for outcome and process measures (see the Appendix). Exclusion of the single study that produced significant outliers (and that reported only process measure data) did not change any of the levels of significance reported, but it did markedly change the reported means for process measures, especially in cells where very few studies were included. To avoid a reporting bias, process measure results are reported both including and excluding the study containing outlying data.

The structure of the data in the present study was such that a proper ANOVA based on a complete matrix of observations was rarely possible. For this reason a number of the reported ANOVAs fail to take into account minor violations of the assumption of independence of observations when two or more observations came from the same study. In such cases, the majority of the studies contributed only one observation. The primary cost of such violations is to inflate the alpha (Stevens, 1992). These situations are noted in the Results and Discussion section.

Where means were compared to each other within the context of an ANOVA, the Newman-Keuls procedure was used. Where mean effect sizes were compared with zero, a Dunnett procedure was used (Glass & Hopkins, 1996). Both procedures use the mean square for the error term of the ANOVA, weighted by the sample size of the cells entering into the comparison. The mean square error allows a more stable estimate of the standard error of the statistic being tested than does the pooled variance of the two cells being compared.

Results and Discussion

Is EMDR Effective?

Outcome measures. When outcome measures were examined, a total of 28 studies with 97 comparisons emerged. We subjected the data to a 7×5 ANOVA with measure type (the five categories of outcome measure described in the Method section) and comparison type (the seven categories of control procedures outlined earlier) as factors. A repeated measures (within-studies) design might seem appropriate here, because some studies used more than one measure and/or comparison type. This yields 35 possible

repeated measures data points for each study, but because many studies had only one comparison group, and no study had more than four, the use of such a design was unworkable, and a between-subjects design was used instead. This is a violation of the independence assumption, but the use of a 1% alpha instead of a 5% alpha should protect against the inflated error rate. For each comparison type we conducted a Dunnett test (no violation of independence) to determine if the mean of the median effect sizes for each study was significantly different from zero. We used Type III sums of squares to deal with the problem of unequal N s.

There was a significant effect of comparison type, $F(6, 67) = 10.22, p < .01$; but no effect of measure type, $F(4, 67) = 1.46, p > .20$; and no significant interaction, $F(19, 67) = 0.63, p > .80$. The mean outcome measure effect sizes by measure type are presented in Table 1. Because there was no difference among these, we collapsed them within studies by taking the median of effect sizes for the outcome measure categories for each study and used these collapsed values in all analyses of outcome data. Again, we followed this procedure so as to not unduly weight studies with measures in more than one category.

These collapsed outcome measure effect sizes by comparison type are presented in Table 2. The effect of comparison type on effect size was significant, $F(6, 44) = 10.53, p < .01$. The largest effect sizes were for comparisons of EMDR versus no treatment, EMDR versus nonspecific treatment, and for pre-post EMDR comparisons within participants. All were significantly greater than zero ($p < .05$).

We examined the same data, but only for studies in which the therapists had been trained by the EMDR Institute. The effect of comparison type on effect size was significant, $F(6, 37) = 10.27, p < .01$. The largest effect sizes were for the same comparison types as in Table 2, and the same comparison types were significantly different from zero. The exclusion of studies using therapists not trained by the EMDR Institute (seven comparisons) makes no difference to the overall statistical comparisons.

Figure 1 shows the collapsed outcome measure effect sizes for the various comparison types for all studies and for those studies in which the EMDR Institute had trained the therapists. We computed 95% confidence intervals for effect sizes using the estimates of standard error from the Dunnett procedure (i.e., using mean square error). The

Table 1
Outcome Measure Effect Size by Measure Type

Measure	Measure type				
	Psychometric PTSD	Psychometric other	Behavioral	Physiological	Other
Mean r	.37 ^a	.32	.28	.04	.39
SD	.43	.36	.42	.43	.40
No. comparisons	28	36	9	8	16
Different from zero (t)	6.37**	6.17**	2.68*	0.35*	5.11**
No. participants required ^b	52	71	95	4,900	46
No. studies required ^c	48	44	80	3,878	36

Note. PTSD = posttraumatic stress disorder.

^a Main effect of difference between effect sizes for various measure types not significant. ^b Participants required to achieve power of 80% with $\alpha < .05$, two tailed, given mean effect size. ^c Studies required to achieve power of 80% with $\alpha < .05$, two tailed, given mean and standard deviation.

* $p < .05$. ** $p < .01$.

Table 2
Collapsed Outcome Measure Effect Sizes by Comparison Type

Measure	Comparison type						
	Pre-post EMDR	No treatment/wait list	Nonspecific treatment	Exposure (not in vivo)	In vivo exposure or CBT	EFixDR	OthDism
<i>r</i>	.64 ^{a,b,c,d}	.44 ^{e,f,g}	.40 ^h	.19 ^a	-.28 ^{b,c,h}	.10 ^{c,f}	.00 ^{d,g}
<i>SD</i>	.31	.25	.30	.19	.31	.24	.16
No. comparisons	14	13	5	3	3	9	4
Different from 0 (<i>t</i>)	10.27**	6.17**	3.43*	1.21	-1.80	1.09	0.00
95% CI	.55 to .72	.31 to .55	.18 to .58	-.12 to .47	-.54 to .02	-.08 to .27	-.26 to .26
No. participants required [†]	14	35	44	212	95 ^{†††}	779	
No. studies required ^{††}	10	12	20	34	44 ^{†††}	190	

Note. Values sharing the same superscript are different ($p < .05$). EMDR = eye movement desensitization and reprocessing; CBT = cognitive-behavioral therapy; EFixDR = EMDR without eye movements; OthDism = other dismantling studies; CI = confidence interval.

[†] Participants in one study required to achieve power of 80% with $\alpha < .05$, two tailed, given mean effect size. ^{††} Studies required to achieve power of 80% with $\alpha < .05$, two tailed, given mean and standard deviation. ^{†††} Because the effect size is negative this is an estimate of the number required to demonstrate an effect for the comparison treatment, not for EMDR.

* $p < .05$. ** $p < .01$.

treatments subsumed as "nonspecific treatments" include: rapid induction and relaxation (Sharpley, Montgomery, & Scalzo, 1996; effect size $r = .74$), biofeedback-assisted relaxation (Carlson, Chemtob, Rusnak, Hedlund, & Muraoka, 1998; effect size $r = .35$), active

listening (Scheck, Schaeffer, & Gillette, 1998; effect size $r = .34$), "routine" individual treatment (Edmond, Rubin, & Wambach, 1999; effect size $r = .18$), and applied muscle relaxation (Vaughan, Armstrong et al., 1994; effect size $r = .32$).

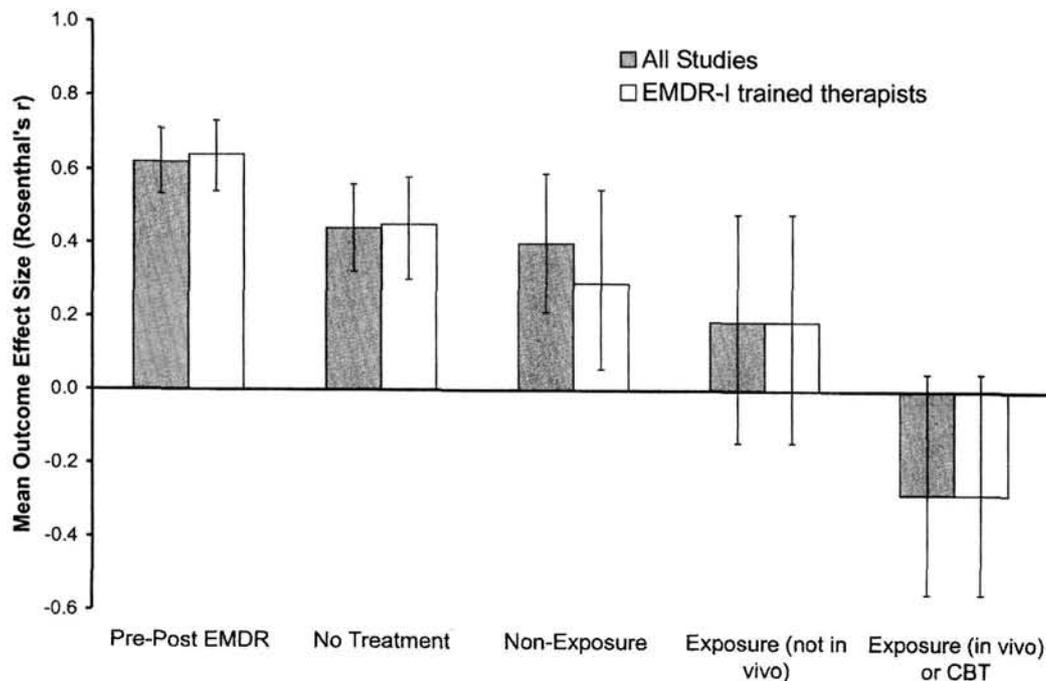


Figure 1. Mean collapsed outcome measure effect sizes (Rosenthal's r) for the comparison of pre-post eye movement desensitization and reprocessing (EMDR within subjects) and of EMDR with other control conditions, both for all studies and for only studies in which therapists were trained by the EMDR Institute. Error bars represent the 95% confidence intervals. A negative effect size indicates that EMDR had a smaller effect size than the comparison condition. Because the distribution of r is asymmetric, values were transformed with Fisher's Z for the purposes of statistical analysis and then re-expressed in terms of r to make the values more intuitively accessible. The confidence interval was calculated in Z score terms and has been re-expressed, like the mean, in terms of r . Note that this produces a confidence interval that is asymmetric about the estimated mean. CBT = cognitive-behavioral therapy.

On outcome measures, EMDR is effective when posttest measures are compared with pretest measures within participants and when EMDR is compared with wait-list/no-treatment or nonspecific treatment controls, regardless of the training of the therapists. It is not shown to be more effective than exposure therapies, EFixDR, or OthDism.

Process measures (SUD and VoC). To examine the effect of EMDR on process measures we performed an ANOVA on the process measures data with comparison type as the factor. Process measures were reported for 18 studies and for six of the seven comparison types (not video or imaginal exposure). Because some of the 18 studies used more than one comparison type, the number of data points is 29. The effect of comparison type on effect size was significant, $F(5, 23) = 5.00, p < .01$. Once again the F test is subject to the violation of independence, so a conservative alpha was used. These data are presented in Table 3. On the process measures, the effect size for EMDR was significantly greater than zero only when pre-post changes within EMDR participants were considered. As might be expected, the effect size for pre-post change within participants was greater than most of the between-groups comparisons (not nonspecific treatment or OthDism), none of which was significantly different from another.

As with the outcome measures, we examined process measures excluding studies in which therapists had not been trained by the EMDR Institute, and again the same pattern of significance existed after these studies were removed.

In summary, on process measures EMDR is effective when posttest measures are compared with pretest measures within participants regardless of the training of the therapists.

Effects of Eye Movements and Alternating Stimuli

Shapiro (1995) stated that

the complete EMDR methodology can afford powerful treatment effects without the use of eye movements or other external stimuli . . . [but that] . . . the addition of such dual attention stimuli increases

the overall speed and efficacy of treatment to a measurable degree. (pp. 327-328)

We examined the assertion that eye movements and other alternating stimuli lead to maximal effect in the current analysis by using the effect sizes arising from comparisons between EMDR and variants of the procedure without eye movements (EFixDR) and dismantling studies between EMDR and variants with other alternating stimuli (OthDism).

EMDR Versus EFixDR. For the EFixDR subanalysis we selected studies in which a clear comparison was made between an eye-movements condition and an eyes-fixed condition, holding the rest of the treatment package essentially constant. One study that was included (Merckelbach, Hogervorst, Kampman, & deJongh, 1994) used a component where the clients engaged in unilateral finger tapping, which is not an alternating movement (Shapiro, 1995). This selection process yielded 13 studies (see Table 4).

Table 5 shows the results of several standard single-sample t tests, on different subsets of the EFixDR data. The process measure effect size was not significantly different from zero regardless of whether the outlier described earlier was included and regardless of whether studies using therapists not trained by EMDR Institute were included. Likewise, the outcome measures effect sizes were not significantly different from zero, regardless of whether all the outcome measures were included individually (that might have added more power, because of the larger number of observations) or they were collapsed (that tended to give all studies equal weight and gave added power by reduced variance within studies).

The published data show no significant incremental benefit because of eye movements. The effect size for process measures (SUD and VoC) is .38 (generally considered medium-large), but this value includes a very large outlier that produces significant heterogeneity, Bartlett test $\chi^2(8, N = 288) = 48.76, p < .001$. Removal of the outlier solves the heterogeneity problem, $\chi^2(7, N = 276) = 8.72, p > .05$. The process measures without the

Table 3
Process Measures Effect Sizes by Comparison Type: All Studies

Measure	Comparison type					
	Pre-post EMDR	No treatment	Nonspecific treatment	In vivo exposure or CBT	EFixDR	OthDism
Mean r	.83 ^{a,b,c}	.47 ^a	.55	-.36 ^b	.15 ^c	.39
SD	.35	.30			.25	.54
Different from 0 (t)	11.63**	2.85			1.19	2.01
No. comparisons	12	4	1	1	8	3
95% CI	.76 to .89	.16 to .70			-.10 to .38	.01 to .68
No. participants required††	7	30	33	54†	343	46
No. studies required†††	4	14			110	68

Note. A statistical outlier was excluded. Values sharing the same superscript are significantly different ($p < .05$). EMDR = eye movement desensitization and reprocessing; CBT = cognitive-behavioral therapy; EFixDR = EMDR without eye movements; OthDism = other dismantling studies; CI = confidence interval. † Because the effect size is negative, this is an estimate of the number required to demonstrate an effect for the comparison treatment, not for EMDR. †† Participants required to achieve power of 80% with $\alpha < .05$, two tailed, given mean effect size. ††† Studies required to achieve power of 80% with $\alpha < .05$, two tailed, given mean and standard deviation.

* $p < .05$. ** $p < .01$.

Table 4
Effect of Eye Movements: Process and Outcome Measure Effect Sizes for Eye Movement Desensitization and Reprocessing (EMDR) and EMDR Without Eye Movements (EFixDR)

Study	No. participants	Median Effect Size (<i>r</i>)	
		Process measures ^a	Outcome measures
Boudewyns et al. (1993)	15	.54	.56
Carrigan & Levis (1999)	36	.16	
Devilly et al. (1998)	24	.16	.16
Dunn et al. (1996)	28	.23	-.09
Feske & Goldstein (1997)	36		.24
Foley & Spates (1995)	20		.06
Gosselin & Mathews (1995)	41	.35	.02
Merckelbach et al. (1994) ^b	39		-.25
Pitman et al. (1996)	34		.06
Renfrey & Spates (1994)	15	-.19	.06
Sanderson & Carpenter (1992) ^b	58	.01	
Tallis & Smith (1994) ^b	24	-.13	
D. L. Wilson et al. (1996) ^c	12	.98	

^a Subjective units of distress and validity-of-cognition ratings. ^b Therapists were not trained by the EMDR Institute. ^c Statistical outlier.

outlier have a mean *r* of .15, and without untrained-therapist studies the mean is .22 (generally considered medium-small). The largest effect size (i.e., including the outlier) would normally be cause for significant findings, but the very large range of the observations (.99 to -.19 with the outlier, and .54 to -.19 without the outlier) produces a very high error term that masks any effect. The effect sizes without the outlier are too small to produce significance. Contrary to the conclusion of Cahill et al. (1999), our analysis indicates that any evidence for the efficacy of eye movements even in the process measures is blurred by the failure to find stable effect sizes across studies.

File drawer effect. The design of this meta-analysis included only published studies. With respect to the specific question on the need for eye movements, one might ask whether unpublished studies exist that will contradict the data presented here. To examine this, we searched Dissertations Abstracts from January 1989 to April 2000 using the key words *EMDR* and *eye movement desensitization*. A total of 46 dissertations met these criteria. The abstracts indicated that 9 of these studies made a comparison between eye-movement and eyes-fixed treatment conditions. Of these, 3 were published and are included in this analysis. The 6 remaining dissertations that compared eye-movement with eyes-fixed conditions all reported no significant difference between the conditions (Johnson, 1996; Lytle, 1993; Marquis, 1995; Onkley, 1993; Opdyke, 1996; and Whalen, 1998). Reviewing Dissertation Abstracts did not yield any unpublished dissertations that showed a significant difference between eye-movement and eyes-fixed conditions. Given the large number of studies required to demonstrate 80% power with the observed mean effect sizes and variance (see Table 5), the available dissertations do nothing to change the picture given by published studies.

Alternating stimuli and OthDism studies. The studies discussed above, in which EMDR was compared with the same procedure without eye movements, are termed *dismantling studies*. OthDism studies have used control techniques involving stimuli or responses on

alternate sides of the body, such as finger tapping using alternate hands or presentation of tones or lights on alternating sides. According to Shapiro (1994, 1995), this sort of alternating stimulus is at least a partial replacement for eye movements. Other dismantling studies have compared EMDR with the same procedure less one component (e.g. Cusack & Spates, 1999). The subanalysis for studies that compared eye movement with eyes fixed is reported above. The subanalysis reported in this section is of the remaining dismantling studies. The current data suggest that the effect sizes of EMDR compared to OthDism on both process and outcome measures are not significantly different from zero. OthDism data were reported in five studies (not including the statistical outlier). The effect size on process measures for EMDR compared to OthDism was .39. Based on three studies, this value is not significantly different from zero, $t(2) = 1.45, p < .05$. The effect size for outcome measures was .00. The very small number of observations makes the EMDR-OthDism comparison a very low-power test. We should treat the relative effectiveness of EMDR to the other dismantling studies (OthDism) as untested by the meta-analysis.

Statistical power for the EFixDR studies. It has been suggested that the studies comparing EFixDR with EMDR have not used sufficient numbers of participants to permit the effect of eye movements to be detected (Cusack & Spates, 1999; Shapiro, 1999). The present data show that the EMDR-EFixDR effect size is marginally significant if one examines only clinical populations satisfying *Diagnostic and Statistical Manual of Mental Disorders* (i.e., *DSM-III*, American Psychiatric Association [APA], 1980; *DSM-III-R*, APA, 1987; *DSM-IV*, APA, 1994) diagnostic criteria (selecting PTSD and Other Anxiety Disorders in Table 6). If we limit ourselves to clinical populations, what sample size would give sufficient power to permit the effect of eye movements to be assessed? We can address this question if all the EFixDR studies are included ($N = 9, r = .10, SD = 0.24$). We can also address it if only groups with *DSM* diagnoses are included ($N = 5, r = .25, SD = 0.24$).

We can look at the sample size needed to find a significant effect using two different methods. One method is to construct a

Table 5
Effect Sizes for Eye Movement Desensitization and Reprocessing (EMDR) Process and Outcome Measures Compared to EMDR Without Eye Movements (EFixDR)

Measure	Effect size (<i>r</i>)			
	Process measures	Process measures without outlier	Outcome measures	Outcome measures collapsed
Mean <i>r</i>	.38	.15	.07	.10
<i>SD</i>	.66	.25	.25	.24
<i>n</i>	9	8	17	9
Minimum	-.19	-.19	-.43	-.25
Maximum	.99	.54	.56	.56
Different from 0 (<i>t</i>)	1.51	1.67	1.13	1.23
No. of participants required ^a	49	343	1,596	779
No. of studies required ^b	128	94	434	190

^a Participants required in one study to achieve power of 80% with $\alpha < .05$, two tailed, given mean effect size. ^b Studies required to achieve power of 80% with $\alpha < .05$, two tailed, given mean and standard deviation.

* $p < .05$.

Table 6
Mean Effect Sizes (and Numbers of Comparisons) for Outcome Measures by Population

Population	Comparison type						
	Pre-post EMDR	No treatment/ wait list	Nonspecific treatment	Exposure (not in vivo)	In vivo exposure or CBT	EFixDR	OthDism
PTSD	.63** (7)	.39* (5)	.34* (2)	.39 (1)	-.44 (1)	.22 (4)	-.06 (1)
Traumatic memories	.58* (4)	.49** (3)	.24* (2)			.00 (3)	.10 (1)
Other anxiety disorders	.74 (1)	.49 (3)		.08 (2)	-.19 (2)	.24 (1)	
Normals	.69 (2)	.47 (1)	.74 (1)			-.25 (1)	-.20 (1)

Note. Numbers in parentheses are numbers of comparisons. EMDR = eye movement desensitization and reprocessing; CBT = cognitive-behavioral therapy. EFixDR = EMDR without eye movements; OthDism = other dismantling studies; PTSD = posttraumatic stress disorder.

* Significantly different from 0 ($p < .05$, Student's t). ** Significantly different from 0 ($p < .00393$: Bonferroni equivalent of .05, Student's t).

single experimental study to test the hypothesis. If the population parameter of r is equal to .10, then a single study with a sample size of 780 in two equal-sized groups (of 390) would be required to achieve a power of 80% to reject the null hypothesis with a Type I error rate of 5% (two tailed). If the population r value is .23, then the required sample size would be 144 (Faul & Erdfelder, 1992).

The other method is to construct a sample of studies in a meta-analytic approach rather than a single experimental study. To do this one needs to take into account the distribution of the effect sizes as well as their means. With a mean effect size of .10 and a standard deviation of .24, we would need 190 studies to show a significant ($\alpha = 5\%$, two tailed) effect. For a mean of .23 with a standard deviation of .24 we would need 38 studies. The meta-analytic numbers seem high because there is some heterogeneity of effect sizes from article to article that is assumed to be absent within a single study. The heterogeneity of effect sizes reflects differences from study to study above and beyond random sampling that cause an increase in the observed variance of effect sizes over the expected value of the variance error of effect sizes estimated from sample size.

Effect Size by Treatment Population

The populations treated with EMDR can be broken into four broad categories. The first category is patients diagnosed with PTSD by one of the *DSM* systems. We recognize that this category has changed considerably from *DSM-III*, to *DSM-III-R*, to *DSM-IV*; however, the studies of EMDR in PTSD have used all three sets of diagnostic criteria. The second population category is people with traumatic memories who do not meet *DSM* criteria. The third category is patients with other anxiety disorders (e.g., specific phobia or panic disorder). The fourth category is normal individuals. Studies in this category attempt to emulate a disorder using normal participants. They generally use a student population and either induce a contrived trauma (e.g., presenting participants with an upsetting photo) and test the treatment, or they use the treatment to decrease the vividness of negative images (e.g., Sharpley et al., 1996).

The data in Table 6 show the effect sizes by population and by comparison type. The number of comparisons in some of the cells is quite small. Inspection of means shows no systematic pattern for population in the data. At the risk of increased Type

I error we conducted 12 one-tailed single-sample t tests on all the means in Table 6 for which the number of comparisons was greater than 1. Consistent with the data pooled across population, the mean effect sizes that were significantly greater than zero occurred only in pre-post comparisons, comparisons with no treatment, and comparisons with nonspecific therapies. Examining the data by population did not reveal anything not seen in the pooled data. Using a more conservative Bonferroni-corrected alpha left the comparison of EMDR with no treatment for participants with traumatic memories and the pre-post comparison within the PTSD population, as significant results.

DeRubeis and Crits-Christoph (1998) argued that EMDR is a potentially effective treatment for patients with noncombat PTSD. There are three studies in this analysis that compared EMDR with treatments not involving exposure, a fair test of its effectiveness (Marcus, Marquis, & Sakai, 1997; Rothbaum, 1997; Vaughan, Armstrong et al., 1994). Comparisons from pre to post EMDR were not included in this section, because such a demonstration does not control for nontreatment effects, a requirement of the criteria set out by Chambless and Hollon (1998). Across the three studies the mean effect size was .45 ($SD = .25$), which is not significantly different from zero, $t(2) = 2.53$, $p < .06$. Given the mean effect size and standard deviation, 18 studies would be required to demonstrate a significant meta-analytic effect with $\alpha = .05$ and a power of 80%. With a point estimate of $r = .45$, a single study with an N of 28 in each treatment group should have a power of 80%. Once again, the need for many studies can be accounted for in terms of between-study variance inflated by very substantial between-study differences in effect size. Although the effect sizes are not significantly heterogeneous, they are nearly so, $\chi^2(2, N = 99) = 5.04$, $p < .09$. Thus, on the basis of the present work there is no clear meta-analytic evidence to support DeRubeis and Crits-Christoph's (1998) assertion that EMDR is a potentially effective treatment in noncombat PTSD, but the evidence suggests that this is a fruitful area for research.

Clinical Significance

One important index of clinical significance is the extent to which posttreatment functioning of treated individuals falls within normal limits on important measures (see Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999). Five studies in the present meta-analysis

addressed this question directly, and the results were consistent with the results of the meta-analysis. Two of the studies assessed clinical significance using normative comparisons as described by Kendall et al. (1999), and the other three used other measures of clinical significance as it is defined in that article. Of the five studies, three compared EMDR with no treatment or with treatment not involving exposure. They all showed the effects of EMDR to be clinically significant. Thus, patients treated with EMDR generally scored in the nonclinical range on relevant instruments, but there was no clinically significant change in the control condition (Feske & Goldstein, 1997; Scheck et al., 1998; S. A. Wilson et al., 1995). The other studies (Devilley & Spence, 1999; Devilly, Spence, & Rapee, 1998) examined EMDR with contrast conditions involving exposure. Both studies found EMDR and the contrast condition to be clinically effective, and no difference in clinical significance was found between the conditions.

Conclusions

The results of this meta-analysis are quite clear for some questions and less clear for others. Does EMDR provide an effective treatment? When outcomes of EMDR treatment are compared with no treatment, and when outcomes are compared with pretreatment status, clients are better off with EMDR treatment than without. The median effect against a comparison of no treatment is strong ($r = .44$, $d = 0.98$)² and reproduced across studies, measures, and complaints. In addition, EMDR seems to be effective compared to nonspecific therapies. This median effect also is strong ($r = .40$, $d = 0.87$).

Is EMDR better than other exposure-based treatments? The answer seems to be no. When outcome measures are considered, EMDR falls into an effect-size category with other treatments that have proven effective, such as exposure treatments for anxiety and cognitive-behavioral therapy for mood. The differences between EMDR and other effective exposure treatments are small enough ($r = .19$, $d = 0.39$ against exposure treatment (not *in vivo*) and $r = -.19$, $d = -0.39$ against *in vivo* exposure or cognitive-behavioral therapy) to make the argument unpersuasive even if we could gather enough data to show significant differences. An exception to this prediction might arise if future research were to show that EMDR has specific utility for a specific patient group or disorder.

Is eye movement—or indeed, any alternating movement—a necessary component of EMDR? The near-zero effect sizes indicate that eye movement is unnecessary. Dismantling studies that examined the use of alternating movements other than eyes were not shown to be different from EMDR by the data available. The effect sizes for the outcome measures are small, and we have no evidence that eye movements or other alternating stimuli are necessary.

Do therapists trained by the EMDR Institute produce different results? The evidence gathered does not show that including only therapists trained by the EMDR Institute changed the conclusions about the effect of EMDR.

Is EMDR more or less effective with some populations than with others? Again there was no solid evidence that this was the case. DeRubeis and Crits-Christoph's (1998) contention that

EMDR is a potentially effective treatment for noncombat PTSD was not supported, although this is perhaps due to the heterogeneity of the findings in the literature.

Early reports portrayed EMDR as a rapid treatment that was generally effective in a single session (e.g., Shapiro, 1989a, 1989b). Much of the basis for these portrayals was process measure data, often in case reports (see Lohr et al., 1992). In the present meta-analysis, within-subject comparisons on process measures (SUD and VoC) do show a spectacular effect size ($r = .81$, $d = 2.71$, based on 12 comparisons). However, we have shown that effect sizes for outcome measures are much more modest. In addition, recent descriptions of EMDR suggest that more than a single session is needed. The treatment manual (Shapiro, 1995) indicates that EMDR is "not one-session therapy" (p. 117), often needing at least 12 sessions (p. 325), and Feske (1998) suggested that the data on EMDR efficacy are clouded because it "was often applied for too few sessions" (p. 178), by which she refers to 3 or fewer sessions. Moreover, Devilly and Spence (1999) showed that 9 sessions of EMDR are less effective than 9 sessions of combined stress inoculation training, prolonged exposure, and cognitive therapy techniques. EMDR is no longer described as a single-session therapy, and the rapidity of EMDR treatment relative to other exposure therapies is not clear from the literature at this time.

What about future research? Some of the articles we read for this project were well written and based on well-designed studies (e.g., Muris, Merckelbach, Holdrinet, & Sijsenaar, 1998, and other work by that group: Cusack & Spates, 1999; Edmond et al., 1999; S. A. Wilson et al., 1995, 1997). Others were not so well designed. The EMDR literature is mature enough now that sloppy studies can be safely left unpublished. The analysis suggests that the noncombat PTSD population is a potentially fruitful area for study. What the literature needs is carefully designed, executed, and analyzed research that pays attention to issues of effect size, power, measurement, and reproducibility as well as issues of clinical significance. The current literature suffers from a substantial variance in findings from study to study, and this may be reduced by improved methodological rigor. Although it makes sense to report SUD and VoC scores for completeness, the use of outcome variables is essential for the assessment of effectiveness. The development of a common battery of outcome measures can progress if researchers make a point of using the best measures found in previous studies along with any new measure they believe to be essential. Such a battery will facilitate normative comparisons for the evaluation of clinical significance across studies as described by Kendall et al. (1999). The progress on dismantling studies is promising, and the responsiveness of the EMDR establishment to the findings of these studies is commendable even if it is frustrating for those trying to track a moving target.

²The notation d refers to Cohen's d , another estimate of effect size. Rosenthal (1991) provided a formula for converting r to d ($d = [2r]/\sqrt{1 - r^2}$), p. 20).

References

References marked with an asterisk indicate studies included in the meta-analysis.

- Acierno, R., Hersen, M., Van Hasselt, V. B., Tremont, G., & Mueser, K. T. (1994). Review of the validation and dissemination of eye-movement desensitization and reprocessing: A scientific and ethical dilemma. *Clinical Psychology Review, 14*, 287-299.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- 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.
- * 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, 10*, 555-569.
- * Bauman, W., & Melnyk, W. T. (1994). A controlled comparison of eye movements and finger tapping in the treatment of test anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 25*, 29-33.
- Boudewyns, P. A., & Hyer, L. A. (1996). Eye movement desensitization and reprocessing (EMDR) as treatment for post-traumatic stress disorder (PTSD). *Clinical Psychology & Psychotherapy, 3*, 185-195.
- * Boudewyns, P. A., Stwertka, S. A., Hyer, L. A., Albrecht, J. W., & Sperr, E. V. (1993). Eye movement desensitization for PTSD of combat: A treatment outcome pilot study. *The Behavior Therapist, 16*, 29-33.
- Cahill, S. P., Carrigan, M. H., & Freuh, B. C. (1999). Does EMDR work? And if so, why?: A critical review of controlled outcome and dismantling research. *Journal of Anxiety Disorders, 13*, 5-33.
- * Carlson, J. G., Chemtob, C. M., Rusnak, K., Hedlund, N. L., & Muraoka, M. Y. (1998). Eye movement desensitization and reprocessing (EMDR) treatment for combat-related posttraumatic stress disorder. *Journal of Traumatic Stress, 11*, 3-24.
- * Carrigan, M. H., & Levis, D. J. (1999). The contributions of eye movements to the efficacy of brief exposure treatment for reducing fear of public speaking. *Journal of Anxiety Disorders, 13*, 101-118.
- Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting & Clinical Psychology, 66*, 7-18.
- * Cusack, K., & Spates, R. (1999). The cognitive dismantling of eye movement desensitization and reprocessing (EMDR) treatment of post-traumatic stress disorder (PTSD). *Journal of Anxiety Disorders, 13*, 87-99.
- DeBell, C., & Jones, R. D. (1997). As good as it seems? A review of EMDR experimental research. *Professional Psychology: Research and Practice, 28*, 153-163.
- DeRubeis, R. J., & Crits-Christoph, P. (1998). Empirically supported individual and group psychological treatments for adult mental disorders. *Journal of Consulting & Clinical Psychology, 66*, 37-52.
- * Devilly, G. J., & Spence, S. H. (1999). The relative efficacy and treatment distress of EMDR and a cognitive-behavior trauma treatment protocol in the amelioration of posttraumatic stress disorder. *Journal of Anxiety Disorders, 13*, 131-157.
- * Devilly, G. J., Spence, S. H., & Rapee, R. M. (1998). Statistical and reliable change with eye movement desensitization and reprocessing: Treating trauma within a veteran population. *Behavior Therapy, 29*, 435-455.
- * Dunn, T. M., Schwartz, M., Hatfield, R. W., & Wiegele, M. (1996). Measuring effectiveness of eye movement desensitization and reprocessing (EMDR) in non-clinical anxiety: A multi-subject, yoked control design. *Journal of Behavior Therapy & Experimental Psychiatry, 27*, 231-239.
- * Edmond, T., Rubin, A., & Wambach, K. G. (1999). The effectiveness of RMDR with adult female survivors of childhood sexual abuse. *Social Work Research, 23*, 103-116.
- Faul, F., & Erdfelder, E. (1992). GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer program]. Department of Psychology, Bonn University, Germany.
- Feske, U. (1998). Eye movement desensitization and reprocessing treatment for posttraumatic stress disorder. *Clinical Psychology: Science & 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.
- * Foley, T., & Spates, C. R. (1995). Eye movement desensitization of public-speaking anxiety: A partial dismantling. *Journal of Behavior Therapy & Experimental Psychiatry, 26*, 321-329.
- * Forbes, D., Creamer, M., & Rycroft, P. (1994). Eye movement desensitization and reprocessing in posttraumatic stress disorder: A pilot study using assessment measures. *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 113-120.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Boston: Allyn & Bacon.
- * Gosselin, P., & Mathews, W. J. (1995). Eye movement desensitization and reprocessing in the treatment of test anxiety: A study of the effects of expectancy and eye movement. *Journal of Behavior Therapy & Experimental Psychiatry, 26*, 331-337.
- * Grainger, R. D., Levin, C., Allen-Byrd, L., Doctor, R. M., & Lee, H. (1997). An empirical evaluation of eye movement desensitization and reprocessing (EMDR) with survivors of a natural disaster. *Journal of Traumatic Stress, 10*, 665-671.
- Greenwald, R. (1996). The information gap in the EMDR controversy. *Professional Psychology: Research and Practice, 27*, 67-72.
- * Hekmat, H., Groth, S., & Rogers, D. (1994). Pain ameliorating effect of eye movement desensitization. *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 121-129.
- * Henry, S. L. (1996). Pathological gambling: Etiologic considerations and treatment efficacy of eye movement desensitization/reprocessing. *Journal of Gambling Studies, 12*, 395-405.
- Herbert, J. D., & Mueser, K. T. (1992). Eye movement desensitization: A critique of the evidence. *Journal of Behavior Therapy & Experimental Psychiatry, 23*, 169-174.
- * Jensen, J. A. (1994). An investigation of eye movement desensitization and reprocessing (EMDR) as a treatment for posttraumatic stress disorder (PTSD) symptoms of Vietnam combat veterans. *Behavior Therapy, 25*, 311-325.
- Johnson, M. D. (1996). *Effects of eye movement desensitization and reprocessing on test anxiety* [CD-ROM]. Abstract from: ProQuest File: Dissertations Abstracts Item: 9716578
- Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Clinical & Consulting Psychology, 67*, 285-299.
- Kleinknecht, R. A. (1993). Rapid treatment of blood and injection phobias with eye movement desensitization. *Journal of Behavior Therapy & Experimental Psychiatry, 24*, 211-217.
- Lipke, H. (1997). Commentary on the Bates et al. report on eye-movement desensitization and reprocessing (EMDR). *Journal of Anxiety Disorders, 11*, 599-606.
- 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 & Experimental Psychiatry, 23*, 159-167.
- Lohr, J. M., Kleinknecht, R. A., Tolin, D. F., & Barrett, R. H. (1995). The empirical status of the clinical application of eye movement desensitization and reprocessing. *Journal of Behavior Therapy & Experimental Psychiatry, 26*, 285-302.

- Lohr, J. M., Lilienfeld, S. O., Tolin, D. F., & Herbert, J. D. (1999). Eye movement desensitization and reprocessing: An analysis of specific versus nonspecific treatment factors. *Journal of Anxiety Disorders, 13*, 185–207.
- Lohr, J. M., Tolin, D. F., & Kleinknecht, R. A. (1996). An intensive design investigation of eye movement desensitization and reprocessing of claustrophobia. *Journal of Anxiety Disorders, 10*, 73–88.
- Lohr, J. M., Tolin, D. F., & Lilienfeld, S. O. (1998). Efficacy of eye movement desensitization and reprocessing: Implications for behavior therapy. *Behavior Therapy, 29*, 123–156.
- Lytle, R. A. (1993). *An investigation of the efficacy of eye-movement desensitization in the treatment of cognitive intrusions related to memories of a past stressful event* [CD-ROM]. Abstract from: ProQuest File: Dissertations Abstracts Item: 9334778
- * Marcus, S. V., Marquis, P., & Sakai, C. (1997). Controlled study of treatment of PTSD using EMDR in an HMO setting. *Psychotherapy, 34*, 307–315.
- Marquis, P. A. (1995). *Eye movement desensitization and reprocessing: Are the eye movements an effective component?* [CD-ROM]. Abstract from: ProQuest File: Dissertations Abstracts Item: 9528198
- McNally, R. J. (1999). EMDR and Mesmerism: A comparative historical analysis. *Journal of Anxiety Disorders, 13*, 225–236.
- * Merckelbach, H., Hogervorst, E., Kampman, M., & deJongh, A. (1994). Effects of "eye movement desensitization" on emotional processing in normal subjects. *Behavioural & Cognitive Psychotherapy, 22*, 331–335.
- * Muris, P., & Merckelbach, H. (1997). Treating spider phobics with eye movement desensitization and reprocessing: A controlled study. *Behavioural & 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 & Clinical Psychology, 66*, 193–198.
- * Muris, P., Merckelbach, H., VanHaften, H., & Mayer, B. (1997). Eye movement desensitization and reprocessing versus exposure *in vivo*. *British Journal of Psychiatry, 171*, 82–86.
- Onkley, P. R. (1993). *A comparison of eye movement desensitization and implosion-like therapy with victims of sexual abuse* [CD-ROM]. Abstract from: ProQuest File: Dissertations Abstracts Item: 9302718
- Opdyke, D. C. (1996). *Clinical efficacy of eye movement desensitization and reprocessing (EMDR) in the treatment of posttraumatic stress disorder* [CD-ROM]. Abstract from: ProQuest File: Dissertations Abstracts Item: 9608510
- * Pitman, R. K., Orr, S. P., Altman, B., Longpre, R. E., Poiré, R. E., & Macklin, M. L. (1996). Emotional processing during eye movement desensitization and reprocessing therapy of Vietnam veterans with chronic posttraumatic stress disorder. *Comprehensive Psychiatry, 37*, 419–429.
- * Renfrey, G., & Spates, C. R. (1994). Eye movement desensitization: A partial dismantling study. *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 231–239.
- Rogers, S., Silver, S. M., Goss, J., Obenchain, J., Willis, A., & Whitney, R. L. (1999). A single session, group study of exposure and eye movement desensitization and reprocessing in treating posttraumatic stress disorder among Vietnam war veterans: Preliminary data. *Journal of Anxiety Disorders, 13*, 119–130.
- Rosen, G. M. (1999). Treatment fidelity and research on eye movement desensitization and reprocessing (EMDR). *Journal of Anxiety Disorders, 13*, 173–184.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.
- * Rothbaum, B. O. (1997). A controlled study of eye movement desensitization and reprocessing in the treatment of posttraumatic stress disorder sexual assault victims. *Bulletin of the Menninger Clinic, 61*, 317–334.
- * Sanderson, A., & Carpenter, R. (1992). Eye movement desensitization versus image confrontation: A single session crossover study of 58 phobic subjects. *Journal of Behavior Therapy & Experimental Psychiatry, 23*, 269–275.
- * Scheck, M. M., Schaeffer, J. A., & Gillette, C. (1998). Brief psychological intervention with traumatized young women: The efficacy of eye movement desensitization and reprocessing. *Journal of Traumatic Stress, 11*, 25–44.
- * Shapiro, F. (1989a). Efficacy of the eye movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress, 2*, 199–223.
- Shapiro, F. (1989b). Eye movement desensitization: A new treatment for post-traumatic stress disorder. *Journal of Behavior Therapy & Experimental Psychiatry, 20*, 211–217.
- Shapiro, F. (1994). Alternative stimuli in the use of EMD(R). *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 89.
- Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols and procedures*. New York: Guilford Press.
- Shapiro, F. (1996a). Errors of context and review of eye movement desensitization and reprocessing research. *Journal of Behavior Therapy & Experimental Psychiatry, 27*, 313–317.
- Shapiro, F. (1996b). Eye movement desensitization and reprocessing (EMDR): Evaluation of controlled PTSD research. *Journal of Behavior Therapy & Experimental Psychiatry, 27*, 209–218.
- Shapiro, F. (1999). Eye movement desensitization and reprocessing (EMDR) and the anxiety disorders: Clinical and research implications of an integrated psychotherapy treatment. *Journal of Anxiety Disorders, 13*, 35–67.
- * Sharpley, C. F., Montgomery, I. M., & Scalzo, L. A. (1996). Comparative efficacy of EMDR and alternative procedures in reducing the vividness of mental images. *Scandinavian Journal of Behaviour Therapy, 25*, 37–42.
- * Silver, S. M., Brooks, A., & Obenchain, J. (1995). Treatment of Vietnam war veterans with PTSD: A comparison of eye movement desensitization and reprocessing, biofeedback, and relaxation training. *Journal of Traumatic Stress, 8*, 337–342.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- * Tallis, F., & Smith, E. (1994). Does rapid eye movement desensitization facilitate emotional processing? *Journal of Behavior Therapy & Experimental Psychiatry, 32*, 459–461.
- * Vaughan, K., Armstrong, M. S., Gold, R., O'Connor, N., Jenneke, W., & Tarrier, N. (1994). A trial of eye movement desensitization compared to image habituation training and applied muscle relaxation in post-traumatic stress disorder. *Journal of Behavior Therapy & Experimental Psychiatry, 25*, 283–291.
- * Vaughan, K., Wiese, M., Gold, R., & Tarrier, N. (1994). Eye movement desensitization: Symptom change in post-traumatic stress disorder. *British Journal of Psychiatry, 164*, 533–541.
- Whalen, J. E. (1998). *EMDR and hypnosis: A theoretical and clinical investigation* [CD-ROM]. Abstract from: ProQuest File: Dissertations Abstracts Item: 9923344
- * Wilson, D. L., Silver, S. M., Covi, W. C., & Foster, S. (1996). Eye movement desensitization and reprocessing: Effectiveness and autonomic correlates. *Journal of Behavior Therapy & Experimental Psychiatry, 27*, 219–229.
- * Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR) treatment for psychologically traumatized individuals. *Journal of Consulting & Clinical Psychology, 63*, 928–937.
- Wilson, S. A., Becker, L. A., & Tinker, R. H. (1997). Fifteen-month follow-up of eye movement desensitization and reprocessing (EMDR) treatment for posttraumatic stress disorder and psychological trauma. *Journal of Consulting & Clinical Psychology, 65*, 1047–1056.

Appendix A
Median Outcome and Process Measure Effect Sizes for Various Comparison Types by Study

Study	Population type	Outcome measure				Process measure				
		Pre-post	Wait list/no treatment	Nonexposure protocol	Video or imaginal exposure or CBT	In vivo exposure or CBT	Wait list/no treatment	Nonexposure protocol	In vivo exposure or CBT	OthDism
Bates et al. (1996) ^{a,b}	Spider phobic	.20								
Bauman & Melnyk (1994) ^{b,c}	Test anxiety		.20							
Boudewyns et al. (1993)	Combat PTSD		.46	.35		.56				.54
Carlson et al. (1998) ^{b,d}	Combat PTSD									.16
Carrigan & Levis (1999)	Speech anxiety		.17			.16				.16
Cusack & Spates (1999)	Trauma memory	.20								
Devilly et al. (1998)	Combat PTSD	.32								
Devilly & Spence (1999)	PTSD									
Dunn et al. (1996)	Trauma memory	.47		.18						
Edmond et al. (1999)	Trauma memory	.42				.24				
Feske & Goldstein (1997) ^d	Panic disorder	.41				.06				
Foley & Spates (1995) ^e	Speech anxiety	.68				.02				.35
Forbes et al. (1994)	PTSD	.65								
Gosselin & Mathews (1995)	Test anxiety	.42								
Grainger et al. (1997)	Trauma memory	.14								
Hekmat et al. (1994) ^{b,d}	Normal students	.84								
Jensen (1994)	Combat PTSD	.30								
Marcus et al. (1997)	PTSD									
Merkelbach et al. (1994) ^{b,d}	Normal students									
Muris et al. (1998) ^d	Spider phobic									
Muris & Merckelbach (1997) ^d	Spider phobic	.74								
Muris et al. (1997)	Spider phobic	.74								
Pitman et al. (1996) ^a	Combat PTSD	.35				.06				
Renfrey & Spates (1994)	PTSD	.74				.06				
Rothbaum (1997) ^d	PTSD									
Sanderson & Carpenter (1992) ^{b,e}	Phobic	.62		.34						
Scheck et al. (1998) ^d	Trauma memory	.85		.74						
Shapiro (1989) ^f	Trauma memory	.70								
Shapley et al. (1996) ^{b,d}	Normal volunteer	.74								
Tallis & Smith (1994) ^{b,c}	Normal	.41								
Vaughan, Weise, et al. (1994) ^d	PTSD (except 2)	.70								
Vaughan, Armstrong, et al. (1994) ^d	PTSD	.74		.32	.40					
D. L. Wilson et al. (1996) ^{e,f}	Trauma memory	.69								
S. A. Wilson et al. (1995)	Trauma memory	.41								

Note. CBT = cognitive-behavioral therapy; EFixDr = eyes fixed desensitization reprocessing; OthDism = other dismantling studies; PTSD = posttraumatic stress disorder.

^a Unable to calculate process measure effect size from data provided. ^b Therapist not trained by Eye Movement and Desensitization Reprocessing Institute. ^c Treatment confound—effect size for outcome measures not calculated. ^d Data on process measures not provided. ^e No outcome comparisons reported. ^f Outlier case.

Received January 11, 2000

Revision received August 3, 2000

Accepted August 4, 2000 ■