COUNTERPOINT: Evaluating EMDR in Treating PTSD

by Shawn P. Cahill, Ph.D.

*Psychiatric Times* ● July 2000 ● Vol. XVII ● Issue 7

(Please see the point discussion by David Serban-Schreiber, M.D., Ph.D.-Ed.)

Eye movement desensitization and reprocessing (EMDR) (Shapiro, 1995, 1989), a relatively new form of psychotherapy, is controversial. I will focus on two aspects of this controversy. The first involves various claims that, compared to other forms of cognitive-behavioral therapy (CBT) that have received empirical support as treatments for posttraumatic stress disorder (PTSD), EMDR is 1) as efficacious, or more so; 2) more efficient, in that it requires fewer sessions; and 3) more acceptable to clients and therapists alike (Pitman et al., 1996; Shapiro, 1999, 1996). The second aspect involves claims that EMDR operates through different (or additional) mechanisms than other forms of CBT, especially exposure therapy (Shapiro, 1999, 1996; Van Etten and Taylor, 1998).

To evaluate the first set of claims, I will summarize the literature on the efficacy of CBT for PTSD and discuss studies in which EMDR has been compared with other forms of treatment. This answers the question, "To what extent have EMDR and CBT been compared in the treatment of PTSD?" Conclusions about the relative merits of any two treatments must be based on direct comparisons in the same randomized study.

To evaluate the second claim, I will summarize dismantling studies that have evaluated the effects of specific components of EMDR. Repeated imaginal exposure to trauma-related memories, an important component of EMDR, has been found to be an effective treatment for PTSD (Foa et al., 1999; Cooper and Clum, 1989; Marks et al., 1998). The principle of parsimony suggests we should assume EMDR affects treatment outcome through imaginal exposure until there is evidence for a contribution of the non-exposure elements of EMDR (e.g., eye movements) to treatment outcome.

Finally, a number of studies evaluating EMDR utilized samples in which participants were not required to meet full *DSM-IV* criteria for PTSD. Of the EMDR studies that compared outcomes of those with full- versus partial-PTSD, the only effect seems to be for diagnosis. Although both groups improved, participants with full-PTSD were more symptomatic than those with partial-PTSD, both before and after treatment (Wilson et al., 1997). There is no evidence for differential improvement as a function of diagnosis (Scheck et al., 1998; Wilson et al., 1997). Therefore, I have chosen to not risk losing potentially valuable information by arbitrarily limiting myself to studies that utilized full-PTSD samples. Instead, I will use the more general term *posttraumatic stress reactions*.
Comparative Efficacy

Cognitive-Behavioral Treatments

There is growing evidence for the effectiveness of three types of (non-EMDR) CBT interventions in the treatment of PTSD. The best-researched is prolonged exposure (PE) (Boudewyns and Hyer, 1990; Foa et al., 1999; Foa et al., 1991).

In PE, clients repeatedly confront thoughts and reminders of the traumatic event until their anxiety decreases, along with their symptoms of PTSD. The treatment is similar to that for other anxiety disorders (Barlow, 1988): intensive, prolonged and repeated imaginal exposure and, in more comprehensive programs, in vivo exposure. For example, the PE protocol developed and evaluated by Foa and colleagues (Foa et al., 1999; Foa and Rothbaum, 1998; Foa et al., 1991) comprises client education, instruction in controlled breathing and seven weekly sessions of imaginal exposure. Daily homework consists of imaginal exposure and gradual in vivo exposure to situations that cause anxiety or avoidance but are objectively safe.

An alternative approach, based on Meichenbaum's work (1977), is stress-inoculation training (SIT) (Foa et al., 1999; Foa et al., 1991). Clients learn a variety of anxiety management skills (e.g., breathing retraining, relaxation, thought stopping, cognitive restructuring and guided self-dialogue) through instruction, role-playing and covert modeling that they then apply in their daily life.

Finally, recent research suggests that variants of cognitive restructuring (Marks et al., 1998; Tarrier et al., 1999) may have promise in the treatment of PTSD. Cognitive restructuring for PTSD helps people identify trauma-related automatic thoughts, evaluate them for accuracy and replace them with more accurate beliefs. Contrary to expectations, combining elements of these different protocols, such as PE with either SIT (Foa et al., 1999) or cognitive restructuring (Marks et al., 1998), has not yielded better outcomes than individual treatments. If anything, adding SIT or cognitive restructuring to PE slightly reduced the short-term efficacy of PE.

Comparisons With Other Treatments

EMDR has been compared with several other treatments utilizing PTSD and PTSR samples, such as relaxation (e.g., Carlson et al., 1998; Vaughan et al., 1994), active listening (Scheck et al., 1998) and "treatment as usual" (Edmond et al., 1999; Marcus et al., 1997). None of these interventions, however, have been independently established to be effective in treating traumatized populations. No published studies have directly compared EMDR with SIT or cognitive restructuring, and only Devilly and Spence (1999) and Vaughan et al. (1994) have directly compared EMDR to a non-EMDR
exposure protocol (Cahill et al., 1999). Further, both of these studies suffer from significant limitations and have yielded contradictory results.

Vaughan et al. (1994) compared EMDR with a treatment called image habituation training (IHT) (Vaughan and Tarrier, 1992). IHT is an imaginal exposure treatment in which clients create six brief tape-recorded descriptions of recurrent intrusive trauma-related images. Each description is followed on the tape by 30 seconds of silence, during which clients imagine the described event as vividly as possible. Clients are instructed to listen to their tape for 60 minutes each day, while recording cognitions and anxiety as homework.

Overall, EMDR and IHT produced comparable results. Shapiro (1996), however, suggested that EMDR was more efficient because IHT required one hour of homework per day in addition to the three to five therapy sessions provided in the study, whereas EMDR did not require homework. Unfortunately, Vaughan et al. (1994) did not report on homework compliance.

Scott and Stradling (1997) raised this concern. In their study, only one of the 14 participants provided with IHT training completed homework as prescribed. Four additional participants completed daily homework of significantly shorter duration, from three to 30 minutes. In the absence of evidence that IHT participants in the Vaughan et al. (1994) study did the prescribed homework, it would be premature to conclude that EMDR was more efficient than IHT because of the homework requirement.

A further limitation of the Vaughan et al. (1994) study is that IHT differs in many ways from the exposure protocols used in other studies of PTSD. Compared to Foa's exposure protocols (Foa et al., 1999; Foa et al., 1991; Marks et al., 1998), IHT utilizes repeated brief exposures to multiple images in rapid sequence, rather than prolonged exposure to a single complete memory followed by focused work on "hot spots." Nor does IHT incorporate in vivo exposure. In addition, there are no independent studies validating the efficacy of IHT for PTSD or comparing it with the more intensive exposure protocols. The only other published outcome study of IHT (Vaughan and Tarrier, 1992) utilized an uncontrolled pretest/posttest design. Scott and Stradling (1997) did not report outcome in their study. Therefore, although participants in the Vaughan and Tarrier (1992) study reported improvement in their symptoms of anxiety and PTSD, the lack of a no-treatment control group precludes attributing improvements to the specific procedures of the intervention.

Devilly and Spence (1999) compared EMDR with a cognitive-behavioral package called Trauma Treatment Protocol (TTP), consisting of prolonged imaginal and in vivo exposure, elements of SIT and additional cognitive therapy interventions. Both protocols called for an initial clinical assessment followed by up to eight treatment sessions. Participants in both groups displayed improvement over the course of treatment, but TTP yielded significantly better outcome than EMDR immediately after treatment and at the three-month follow-up assessment. These authors also devised a self-report mea-sure to
assess distress levels and intrusiveness produced by the treatments and found them to be equivalent on these variables.

Although these results suggest that this comprehensive cognitive-behavioral intervention is superior to EMDR, they need to be interpreted with caution. The study's primary methodological weakness is that participants were not randomly assigned to conditions. Instead, the first 20 participants were treated in two groups of 10, with the first 10 referrals treated with one intervention and the second 10 treated with the other intervention, although it had been determined by chance that TTP would be administered first. Additional participants were randomly assigned to treatment conditions. The final participant who was supposed to be assigned to EMDR received TTP instead, because preliminary analyses already indicated better outcomes with TTP. The authors, therefore, felt ethically obliged to provide this participant with TTP.

Unfortunately, re-assigning participants on the basis of preliminary analyses assumes the very outcome the study was designed to evaluate, possibly biasing results in favor of the nominally superior treatment. A second limitation of this study is that the senior author, who synthesized the TTP intervention, served as the primary therapist in both treatment conditions. This raises the possibility of differential alliance to or familiarity with the two treatments as another alternative explanation for the findings.

In summary, neither the study by Vaughan et al. (1994) nor the one by Devilly and Spence (1999) provides an adequate basis for determining the relative efficacy, efficiency or acceptability of EMDR and CBT. Given the absence of any published, methodologically sound studies that directly compare EMDR and CBT, there is presently no adequate empirical basis for drawing conclusions about the relative merits of EMDR and CBT.

**Mechanisms of EMDR**

Dismantling studies identify important elements of a treatment package by comparing the full treatment with variations in which one or more components have been removed. These studies of EMDR focus mainly on eye movements or other laterally alternating stimuli. A recent narrative review (Cahill et al., 1999) identified six studies utilizing PTSD populations that compared EMDR with a no-eye-movement condition. Four of the six studies in Cahill et al.'s review (1999) found no differences between the two conditions. The two remaining studies found EMDR superior to the no-eye-movement condition on assessment measures taken during treatment sessions, but not on posttreatment outcome measures.

For example, both studies found EMDR produced greater reductions in subjective units of distress (SUD) ratings obtained during treatment sessions. Boudewyns et al. (1993) also rated more participants in the EMDR condition as treatment responders than in the no-eye-movement condition. Analysis of skin conductance levels during the first and last treatment sets in the Wilson et al. study (1996) showed within-set reduction for the EMDR group, in contrast to no within-set changes for the control group. However, no
group differences were found on any outcome measure in the Boudewyns et al. (1993) study, not even posttreatment SUDs.

Neither group improved significantly, nor were they any different from a third standard care control group. Wilson et al. (1996) did not assess treatment outcome before administering EMDR to all but one of the control participants. Thus, while there is some evidence from studies with PTSD (Boudewyns et al., 1993) and PTSR (Wilson et al., 1996) samples that eye movements may have some effect on within-session measures of anxiety, there is no evidence that eye movements improve treatment outcome.

Some proponents of EMDR have questioned the validity of conclusions drawn from many of the group dismantling studies cited above. This has been for reasons related to inadequate treatment fidelity, such as too few sessions for the population. For example, Boudewyns et al. (1993) and Devilly et al. (1998) both treated veterans utilizing only two sessions. A discussion of treatment fidelity by proponents of EMDR can be found in Greenwald (1996), Lipke (1999) and Shapiro (1996). Rosen (1999) offers another alternative perspective. It should be noted, however, that simply identifying limitations of existing research does not justify assuming the results would necessarily have been different, had the researchers just "done it right."

Montgomery and Ayllon (1994) conducted a dismantling study utilizing a complex multiple-baseline design across three pairs of individuals diagnosed with PTSD. The study consisted of four phases in which an initial baseline phase (phase A) was followed by a treatment phase (phase B) that included all EMDR components except the eye movements. The third phase (phase BC) consisted of adding eye movements to the intervention, after which there was a follow-up period that was procedurally identical to the baseline phase. The dependent variables in this study were the Beck Depression Inventory (BDI), obtained at the beginning and end of the study; weekly reports of the number of days with intrusive thoughts and disturbing dreams, averaged across each phase; mean heart-rate and systolic blood pressure obtained during each session, averaged across phases; and mean SUD ratings obtained during sessions.

Before discussing the results of this study, it may be useful to review the criteria by which multiple-baseline studies across participants are evaluated. More is required to make causal inferences than just showing a change on the dependent variable following the phase shift that is replicated in multiple individuals (Barlow and Hersen, 1984). The multiple-baseline design across participants begins with obtaining concurrent baselines on multiple individuals. Once stable baselines are obtained for all participants, the experimental treatment is introduced to one participant while the baseline conditions are maintained for the remaining participant(s). When the target individual's response during the experimental phase has stabilized, the experimental manipulation is then introduced to the next individual assuming that concurrent response to the baseline condition has remained stable for the other participant(s). This process is repeated until all participants have received the intervention. The intervention's effectiveness is demonstrated when the dependent variable changes with the phase shift for the treated participant, but not the untreated participant(s). This pattern is subsequently replicated across participants.
A logical prerequisite to meeting these conditions is that, for each person, multiple data points within each phase must be available for visual inspection. Unfortunately, this requirement is not met for the majority of measures in the Montgomery and Ayllon (1994) paper. In fact, it is only met for SUD levels. Casual inspection of the relevant graphs suggests that substantial decreases in SUD ratings occurred during the B (no-eye-movement) phase for only one of the six participants (subject 4, and then only in the first B session). By contrast, substantial decreases in SUD scores occurred during the BC (full EMDR) phase for five of the six participants (all but subject 5). These observations may appear to support the hypothesis that eye movements enhanced fear reduction. A more careful inspection of some of the participant pairs, however, cautions against concluding that eye movements were responsible for the decline in SUD ratings during the BC phase.

In the first pair of participants, subject 1 was shifted from the B to BC phase between session 7 and session 8, with little difference between SUD levels on these two days. Although subject 1's SUD ratings subsequently declined over the course of the BC phase, the decline also continued throughout the follow-up phase. In other words, this subject's SUD levels never become stable in the BC phase. Nevertheless, subject 2 was shifted from the B to BC phase between session 8 and session 9. This shift of subject 2 only one session after shifting subject 1 precludes comparing the decline observed in the BC phase for subject 1 with an ongoing B phase for subject 2. Thus, we cannot confidently attribute the decline in SUD ratings for both individuals from session 9 to session 13 to the eye movements, as there is no concurrent no-eye-movement condition against which to compare the eye-movement condition.

With regard to the third pair of subjects, the interpretive problem here is that SUD ratings for one of the two individuals (subject 5) did not show much decline during the BC phase, while the other one (subject 6) did. This lack of consistency across the two participants raises questions as to whether the decline in subject 6's SUDs can actually be attributed to the eye movements.

Thus, there is no solid evidence in the Montgomery and Ayllon (1994) study that meets the criteria for drawing causal inferences from multiple-baseline designs to support conclusions about the importance of eye movements in EMDR. Further, the unavailability of correspondingly fine-grained data for the BDI scores, weekly symptom reports and psychophysiological measures prevents conclusions about whether the observed changes in symptom measures over time can be attributed to any specific component of the intervention.

A recent study by Cusack and Spates (1999) is the only one to investigate the role of "installation" trials—the cognitive restructuring component of EMDR—for PTSR. Participants were randomly assigned to receive either standard EMDR or a condition in which the installation trials were replaced by additional desensitization trials. Both groups improved during the study and retained their improvements at two-month follow-up. There were no differences between groups on any measure. Thus, as with the eye movements, there is no evidence that the other major non-exposure element of EMDR-its unique form of cognitive restructuring-improves treatment outcome.
Discussion

I have attempted to illustrate that the primary literature on EMDR does not justify claims about its relative efficacy, efficiency and acceptability in comparison to CBT. Nor is there any strong evidence that EMDR achieves its therapeutic effects through different or additional mechanisms than exposure therapy. If the primary literature does not support such claims, then where are they from?

Many are based on authors making informal comparisons across studies (Lipke, 1999; Montgomery and Ayllon, 1994; Pitman et al., 1996). Given the often substantial differences across various studies of EMDR and CBT (e.g., different samples, different measures, single versus multiple therapists, differing duration and number of sessions, different control groups, and so on), such comparisons are fraught with difficulties (Cahill and Frueh, 1997) and do not provide an adequate basis for drawing conclusions about comparisons between treatments.

A second basis for such assertions is the use of meta-analysis, a procedure intended to provide a quantitative method for reviewing and synthesizing the results of research studies. One recent comprehensive meta-analysis of treatments for PTSD concluded:

> Behaviour therapy and EMDR were the most effective psychological therapies, and both were as effective as SSRIs [selective serotonin reuptake inhibitors]. Effect sizes were large across all PTSD symptom domains for these treatments in relation to controls, and treatments were generally statistically comparable in efficacy (Van Etten and Taylor, 1998).

The authors further suggested that EMDR is more efficient than other treatments, and that EMDR achieves its therapeutic effect through some mechanism other than exposure.

The studies included in their meta-analysis, however, did not include a single study in which EMDR was directly compared with behavior therapy which, as they defined it, combined studies of PE, SIT and IHT. The Vaughan et al. study (1994) was not included because only 80% of participants met criteria for PTSD (S. Taylor, personal communication, January 1999) and the Devilly and Spence (1999) study had not yet been published. Nor is there a single study in their meta-analysis in which any form of psychotherapy was directly compared with medication.

Conclusions drawn from meta-analysis are heavily dependent on the methods used to identify and select studies, compute the effect sizes, and group the various studies. In order to increase the number of studies included in their meta-analysis and create comparisons across studies that do not exist in the primary literature, Van Etten and Taylor (1998) did not use the standard method for computing effect sizes.

The standard method is to compute a between-group effect size by subtracting the posttreatment mean of the comparison group-of-interest from the corresponding mean of
the target treatment group, and then dividing this group difference by the pooled standard deviation (Cohen, 1988). This is done for all comparisons-of-interest in each study to be included in the meta-analysis. The resulting effect sizes from the different studies are then combined according to the types of comparisons of interest.

In contrast, Van Etten and Taylor (1998) computed within-group effect sizes. For each group-of-interest in a study, the posttreatment mean was subtracted from the pretreatment mean and divided by the pooled within-group standard deviation. Their rationale was that this allowed inclusion of uncontrolled studies in their meta-analysis, as a control group is not necessary for computing within-group effect sizes, thereby "increasing the number of trials and statistical power to detect differences between treatments." They subsequently categorized the within-group effect sizes in terms of the type of intervention and compared average effect sizes across categories. It is important to understand that there was no overlap in studies between the 13 effect sizes for behavior therapy and the 11 effect sizes for EMDR in the Van Etten and Taylor (1998) meta-analysis.

There is a serious concern with using within-group effect sizes to create comparisons across groups of studies that do not exist in the primary literature. It ignores that study populations are necessarily nested within their studies and that, in the absence of direct comparisons between therapy types, the different studies are themselves nested within their type of treatment. This confounds study samples with type of treatment, precluding meaningful conclusions about the comparative efficacy of the different treatments.

Consider the comparison of average effect sizes on total self-reported PTSD severity between behavior therapy and EMDR (1.27 and 1.24, respectively). These mean values tell us that, within each set of studies, the average within-group effect sizes for the different treatments were quite similar. They do not say, however, what the average effect size of EMDR would have been in the populations represented in the behavior therapy studies, nor do they specify what the average effect size of behavior therapy would have been in the populations represented in the EMDR studies. Furthermore, there is no basis for assuming that, just because all of the studies in the meta-analysis utilized full PTSD samples, that there would be no differences across the various study samples in such variables as severity, chronicity or motivation for (and responsiveness to) treatment.

The danger of this strategy in Van Etten and Taylor's (1998) meta-analysis may be further illustrated by considering an example in which it yields a conclusion different from the one drawn from the relevant primary literature. The meta-analysis included only one effect size for an EMDR group without eye movements.

They noted, "When all eye movement conditions in the meta-analysis were compared with this one fixed-eye condition, EMDR was more effective. However, when the fixed-eye control was compared to the EMDR condition within the same study [i.e., Devilly and Spence, 1996], the fixed-eye condition was comparable to the EMDR condition." (The Devilly and Spence study cited by Van Etten and Taylor was an unpublished manuscript at that time. It was later published as Devilly et al., 1998.)
Furthermore, I have already mentioned numerous dismantling studies not included in Van Etten and Taylor's meta-analysis that concluded that eye movements did not contribute to treatment outcome (Cahill et al., 1999).

In the absence of any convincing evidence from dismantling studies that any of the unique features of EMDR contribute to treatment outcome, the remaining basis for claims that EMDR operates through mechanisms other than (or in addition to) exposure is a logical argument (Shapiro, 1999, 1996; Van Etten and Taylor, 1998). Proponents argue that the amount of exposure in the EMDR protocol is less than in exposure-therapy protocols and is implemented in ways that are less than optimal for exposure therapy (e.g., brief interrupted exposures in EMDR versus long, uninterrupted exposures in PE). Since EMDR achieves the same or better outcome as PE in the same or fewer sessions, exposure alone cannot be the operative mechanism. One hopes it can be seen that this argument rests on an unsubstantiated assumption about the relative efficacy/efficiency of EMDR and PE. As such, the conclusion is uncertain.

Questions remain as to the crucial components of effective treatments and their relative merits. Narrative reviews and meta-analyses are useful means of summarizing accumulated knowledge and generating hypotheses. Logical analyses are also helpful in generating new hypotheses and for guiding new studies. None of these methods, however, replace sound empirical research as the primary basis for the growth of scientific knowledge.

Dr. Cahill is an instructor at the Center for the Treatment and Study of Anxiety in the department of psychiatry at the University of Pennsylvania School of Medicine. He gratefully acknowledges Steven Taylor, Ph.D., for providing information regarding studies excluded from the Van Etten and Taylor (1998) meta-analysis.

References


Retrieved www 8/31/05 http://www.psychiatristtimes.com/p000741.html