Prior meta-analyses have suggested that eye-movement desensitization and reprocessing (EMDR) may be effective in alleviating the symptoms of post-traumatic stress disorder (PTSD). EMDR is now being recommended as a treatment for military combat veterans who suffer from PTSD. We provide a review of published outcome studies that appeared in print from 1987 – April, 2008 which examined the specific effects of EMDR on PTSD among military combat veterans. Studies were identified through electronic bibliographic databases, web sites, and manual searches of article reference lists. A total of six randomized controlled trials (RCTs) and three quasi-experimental studies met our inclusionary criteria and are reviewed. The evidence supporting the use of EMDR to treat combat veterans suffering from PTSD is sparse and equivocal, and does not rise to the threshold of labeling the therapy as an empirically supported treatment. It is premature to incorporate EMDR into routine care for veterans to alleviate combat-related PTSD. EMDR needs a considerably stronger evidentiary foundation which includes large-scale RCTs involving credible placebo controlled treatment conditions. Copyright © 2009 John Wiley & Sons, Ltd.

The United States military regularly places people in settings that present a high risk for exposure to violent trauma. More than 75% of US military forces stationed in Afghanistan and Iraq report experiencing events that involved threatened death or serious injury (Mental Health Assessment Team, 2006). Trauma exposure has been causally linked with the development of physiological diseases (Boscarino, 1997) and psychiatric disorders (Hoge, Castro, Messer, McGurk, Cotting, & Koffman, 2004; Jordan, Schlenger, Kulka, Weiss, Fairbank, & Marmar, 1991). Exposure to trauma is a principle risk factor for mental health problems in combat settings (Fontana & Rosenheck, 1998).

The Department of Veteran Affairs (2008) reported that since 2002 it has diagnosed over 120,049 (40%) of the Iraq and Afghanistan veterans with a possible mental health disorder and approximately 60,000 veterans with a probable diagnosis of post-traumatic stress disorder (PTSD). PTSD is an Axis I anxiety disorder that can develop after exposure to a traumatic event or experience, including those experienced during
combat. It is correlated with experiencing military combat situations (Hoge, Auchterlonie, & Milliken, 2006; Kulka, Schlerenger, Fairbank, Hough, Jordan, Marmar, et al., 1990) as well as other unusual traumatic events (e.g., sexual assault, natural disasters, accidents, etc.).

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM, American Psychiatric Association, 2000), in order to be diagnosed with PTSD, a person must have experienced, witnessed, or confronted death or serious bodily injury to self or other and responded with intense fear, helplessness, or horror. Symptoms are said to appear in three clusters, i.e., re-experiencing, avoidance/numbing, and hyperarousal; must last greater than 1 month; and cause significant clinical distress or impairment in social, occupational, or other functioning (American Psychiatric Association, 2000).

Eye-movement desensitization and reprocessing (EMDR) was introduced as a cognitive-behavioral therapy for clients with anxiety disorders almost 20 years ago (Shapiro, 1989). During a treatment session the client is instructed to focus on recalling a traumatic memory while simultaneously focusing on an external stimulus such as tracking the therapist’s finger being waved back and forth in front of the client in a precisely prescribed manner. EMDR purports to access and process traumatic memories and facilitate the desensitization of emotional distress, the reformulation of associated cognitions, and relief of accompanying physiological arousal (Shapiro, 1995, 1999).

The treatment of PTSD with EMDR has been studied extensively. Previous meta-analyses (Van Etten & Taylor, 1998; Alto, 2001; Davidson & Parker, 2001; Maxfield & Hyer, 2002; Bisson & Andrew, 2005; Bradley, Greene, Russ, Dutra, & Westen, 2005; Seidler and Wagner; Bisson, Ehlers, Matthews, Pilling, Richards, & Turner, 2007), and various reviews (Rubin, 2003, 2004; Maxfield, Lake, & Hyer, 2004; Shepherd, Stein, & Milne, 2000; Hertlein & Ricci, 2004; Institute of Medicine, 2007) all examined the provision of some form of EMDR on PTSD among diverse populations and settings ranging from adult female survivors of childhood sexual abuse (Edmond, Rubin, & Wambach, 1999) to an HMO setting (Marcus, Marquis, & Sakai, 1997). The general conclusion is that EMDR does produce moderate improvements in PTSD symptomatology. Most comparison groups involved waiting lists condition or usual care. Randomized controlled trials (RCTs) examining EMDR head-to-head with other accepted treatments such as gradual real-life exposure therapy or documented credible placebo control conditions are rare.

Given EMDR’s apparent usefulness with PTSD as a general diagnostic category, various groups are now recommending its use to treat PTSD among military combat veterans (e.g., EMDR Institute, Department of Veterans Affairs, 2004; Russell, Silver, Rogers, & Darnell, 2007), and it is currently being applied as a therapy option within the military (Lewis, 2009). The etiological precipitants of PTSD can vary widely,
involving experiences as diverse as being sexually assaulted, surviving a natural disaster, being in an automobile accident or airplane crash, or being exposed to military combat. There is some preliminary evidence that combat-related PTSD differs significantly from other forms of psychopathology (e.g., Koenen et al., 2002, 2003; Fu et al., 2007) and questions have been raised regarding the external validity of studies on psychosocial treatments for PTSD in general (encompassing a wide array of etiological pathways) to samples of military veterans whose PTSD stems from combat exposure in particular (Stirman, 2008). As noted by Keane, Silberbogen and Weierich, (2008, p. 304), the traumatic events causing PTSD in a given individual ‘...may vary on dimensions that are of particular salience in case conceptualization and treatment planning’. Prior reviews of the EMDR literature have not isolated the evidence specifically addressing the usefulness of this approach in the treatment of the combat veteran with a diagnosis of PTSD. To help fill this lacunae, we completed such a review in order to answer the following question: Does EMDR reduce PTSD symptomatology in military combat veterans?

METHOD

Selection of Studies

Studies were identified through electronic bibliographic databases, Web sites, and manual searches. Databases searched included ACP Journal Club, ASSIA: Applied Social Sciences Index and Abstracts, CINAHL with Full Text, Cochrane Database of Systematic Reviews, DARE, Dissertation Abstracts/Digital Dissertations, MEDLINE, PILOTS, PsychINFO and Social Services Abstracts. Web sites, such as the Agency for Healthcare Research and Quality (www.ahrq.gov), Department of Veteran Affairs (www.va.gov), ClinicalTrial.gov (www.clinicaltrial.gov), EMDR Institute, Inc. (http://www.emdr.com), the National Institute of Mental health (www.nimh.nih.gov), and David Baldwin’s Trauma Pages (http://www.trauma-pages.com), were searched. Manual searches of the reference sections of identified articles, government reports, meta-analyses, pertinent published books, and systematic and non-systematic reviews were also conducted. Our review did not undertake a search for unpublished data, hence this cannot be considered to be a true systematic review. However, attempts were made to contact known experts to identify any additional published reports. The full search terms used to guide the database searches included the following: ‘EMDR’ or ‘eye movement desensitization and reprocessing’ AND ‘PTSD’ or ‘post-traumatic stress disorder’ AND ‘veteran’; or ‘soldier’ or ‘military personnel’ or ‘combat’ or ‘war’ or ‘World War I’ or ‘World War II’ or ‘Korean War’ or ‘Korean Conflict’ or ‘Vietnam War’ or ‘Gulf War’ or ‘Operation Enduring Freedom’ or ‘OEF’ or ‘Operation Iraqi Freedom’ or ‘OIF’.
Search terms were modified to meet the requirements of individual databases in regard to differences in fields. Located studies were chosen for initial review if they met the following criteria: (a) randomized controlled trial or quasi-experiment, (b) PTSD diagnosis based on DSM criteria, (c) PTSD outcome measure, (d) Published between 1987 (the inception of EMDR) and April 2008, (e) English language. Using this approach, 15 studies were located, of which three were case studies (Young, 1995; Carlson, Chemtob, Rusnak, & Hedlund, 1996; Russell, 2006), one was a clinical comment (Thomas & Gafner, 1993), one discussed training needs (Russell & Silver, 2007) and another did not otherwise meet our inclusionary criteria. A total of six RCTs and three quasi-experimental studies were retained for review and are discussed below.

Analytical Strategies

Hertlein and Ricci’s (2004) Platinum Standard (PS) Scale was used to evaluate studies’ methodological characteristics. This instrument expands upon Maxfield and Hyer’s (2002) Revised Gold Standard (RGS) Scale and Foa and Meadows (1997) Gold Standard (GS) Scale. It assesses a study’s methodological quality across 13 standards (e.g., clearly defined target symptoms, assessor reliability, treatment adherence, length of treatment, effect size reporting). Each of the individual 13 items are scored by a rater as 0, 0.5, or 1, with a range of possible total scores for the overall scale from 0 (none of the 13 standards was met) to 13 (all 13 of the standards were adequately met).

A peer rater, who was blind to the question of the review, worked independently from the first author. Both the first author and external rater evaluated each of the nine studies across the thirteen standards of the PS Scale. Of the 234 ratings (13 ratings per article × nine articles × two raters) completed by the two raters, only nine (4%) were in disagreement; that is, 96% were in agreement. In these cases, differences were identified by the two raters and were resolved through consensus. The total PS score for each study ranged from 4 to 9.5, with a mean of 8 (SD = 1). This is above the median (7) score indicating a slightly above average assessment of reviewed studies.

RESULTS

Study Characteristics

Rusnak, Hedlund, and Muraoka (1998), Devilly, Spence, & Rapee (1998), Rogers, Silver, Goss, Obenchain, Willis, and Whitney (1999), and Macklin, Metzger, Lasko, Berry, Orr, and Pitman (2000). All studies reviewed recruited their participants from Veteran Administration-related facilities and/or clinics. An overview of the nine articles is shown in Table 1.

Various instruments were used to measure initial client functioning and treatment outcomes. However, only outcome measures of PTSD are presented in this review divided into the categories of self-report instruments and structured interviews. Self-report instruments included the Impact of Event Scale (IES) (Horowitz, Wilner, & Alvarez, 1979) and the Mississippi Scale for Combat-Related Post-Traumatic Stress Disorder (M-PTSD) (Keane, Caddell, & Taylor, 1988; McFall, Smith, & Mackay, 1990). Structured interviews included the Clinician-Administered Post-Traumatic Stress Disorder Scale (CAPS) (Blake, Weathers, Nagy, Kaloupek, Charney, & Keane, 1990), the Post-Traumatic Stress Disorder Interview (PTSD-I) (Watson, Juba, Manifold, et al., 1991), the Structured Clinical Interview (SCID) (Spitzer, Williams, & Gibbon, 1987), and the Structured Interview for Post-Traumatic Stress Disorder (SI-PTSD) (Davidson, Smith, & Kudler, 1989).

**Study Outcomes**

The Boudewyns et al. (1993) study utilized a sample of 20 Vietnam combat veterans diagnosed with PTSD as conceptualized in the DSM-III-R (American Psychiatric Association, 1980) and measured by the CAPS. They were randomly assigned to one of three groups: EMDR (n = 9), Exposure (n = 6), or Control (n = 5). The therapist was personally trained by Francine Shapiro, the founder of EMDR, though to what level is unclear. As reported in Alto (2001), a later research review (Lohr et al., 1999) reported that the assessors were not blind to treatment conditions. The treatment was apparently replicable and specific but not consistent with the standard EMDR protocol given the limited number of sessions (only two). Treatment adherence was unclear as was whether or not there were non-confounded conditions. The instruments used to assess post-treatment outcomes were the CAPS, the IES, and the M-PTSD. It is unclear whether or not the assessor was blinded and received training to administer the instruments. The authors reported that ‘a 2 (treatment condition) by 2 (pre- to post-therapy) repeated measures ANOVA yielded no significant effects for any of the psychological measures’ (Boudewyns et al., 1993, p. 31); however, no data were actually provided to corroborate this finding. Significant limitations to this study included lack of detail regarding the assessor, treatment adherence and qualifications for exposure condition, and lack of statistical power.
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Outcome measure</th>
<th>Effect</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Boudewyns, Stwertka, Hyer, Albrecht, and Sperr, 1993</td>
<td>EMDR ($n = 9$), Exposure ($n = 6$), Control ($n = 5$)</td>
<td>CAPS, IES, M-PTSD</td>
<td>No improvement seen on psychological measures</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; Random assignment; Multiple outcome measures; No-treatment control group</td>
<td>Use of blind, independent assessor unclear; Assessor reliability unclear; Treatment adherence unclear; Few participants exposed to confounds with no control for variables; 2 sessions; Qualifications not provided for comparative group, clinicians; Low statistical power</td>
</tr>
<tr>
<td>2. Jensen, 1994</td>
<td>EMDR ($n = 13$), Control ($n = 12$)</td>
<td>M-PTSD, SI-PTSD</td>
<td>Failed to support effectiveness of EMDR</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; Random assignment; Multiple outcome measures; No-treatment control group</td>
<td>Use of blind, independent assessor unclear; Assessor reliability unclear; Treatment adherence contentious; Nonconfounded conditions unclear; 1 intake and 2 treatment sessions; Low statistical power</td>
</tr>
<tr>
<td>3. Silver, Brooks, &amp; Obenchain, 1995</td>
<td>EMDR ($n = 13$), Relaxation ($n = 9$), Biofeedback ($n = 6$), Control ($n = 55$)</td>
<td>Problem Report Form (PRF), no single measure, 8 symptom scales</td>
<td>Effective on non-standardized (PRF) measure</td>
<td>DSM diagnosis; No-treatment control group</td>
<td>Non-standardized measure; Use of blind, independent assessor not reported; Assessor reliability not reported; Not standard EMDR protocol; Assignment not randomized; Treatment adherence not reported; Most participants exposed to confounds with no control for variable; Self-report measures only; Length of treatment variable, but less than six sessions; Level of therapist training unclear; Low statistical power</td>
</tr>
<tr>
<td>Study (Authors)</td>
<td>Treatment Groups</td>
<td>Outcomes Measures</td>
<td>Findings</td>
<td>Methodological Considerations</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>----------</td>
<td>-------------------------------</td>
<td></td>
</tr>
<tr>
<td>Boudewyns &amp; Hyer, 1996</td>
<td>EMDR (n = 21),</td>
<td>CAPS, IES</td>
<td>CAPS showed decrease in all 3 conditions; IES showed no significant changes</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; Assessor was blind; Random assignment; Multiple outcome measures; No-treatment control group. Assessor reliability unclear; Treatment adherence unclear; 5-7 treatment sessions; Qualifications for treatment group, clinician provided but unclear on level of EMDR training; low statistical power.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exposure (n = 18), Control (n = 22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitman, Orr, Altman, Longpre, Poire, &amp; Macklin, 1996</td>
<td>Eye movement (n = 7), Eyes fixed (n = 10)</td>
<td>CAPS, IES, M-PTSD</td>
<td>Moderate improvement in both groups</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; Assessor was blind and independent; Random assignment; Treatment fidelity independently checked and adequate; Multiple measures; Qualifications for treatment and comparative group, clinicians provided. Not standard EMDR protocol; Unclear adherence; Less than 6 treatments; EMDR condition attrition; Low statistical power.</td>
<td></td>
</tr>
<tr>
<td>Carlson, Chemtob, Rusnak, Hedlund, &amp; Muraoka, 1998</td>
<td>EMDR (n = 10), Biofeedback (n = 13), Routine Care (n = 12)</td>
<td>CAPS, IES, M-PTSD, PTSD Symptoms Scale</td>
<td>Significant treatment effects in EMDR condition</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; Random assignment; Multiple outcome measures; 12 sessions; No-treatment control group. Assessor was not blind; Assessor reliability unclear; Treatment adherence variable; Nonconfounded conditions unclear; Therapist trained, but level unclear; Low statistical power.</td>
<td></td>
</tr>
<tr>
<td>Devilly, Spence, &amp; Rapee, 1998</td>
<td>51</td>
<td>M-PTSD</td>
<td>No statistically significant improvement in effects</td>
<td>DSM diagnosis; Reliable, valid, and adequate measure used; Qualifications for treatment and comparative group, clinicians provided; No-treatment control group. Assessor was a therapist; Assessor reliability was unclear; Not standard EMDR protocol; Only one therapist; No fidelity checks; Multiple measures; 2 sessions; low statistical power.</td>
<td></td>
</tr>
</tbody>
</table>

(Continues)
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Outcome measure</th>
<th>Effect</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Rogers, Silver, Goss, Obenchain, Willis, &amp; Whitney, 1999</td>
<td>EMDR (n = 6), Relaxation (n = 6)</td>
<td>CAPS, IES</td>
<td>Both conditions improved in intrusive symptoms; No improvement in avoidance symptoms</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; Assessor was blind; Random assignment; Multiple measures</td>
<td>Not standard EMDR protocol; Assessor reliability unclear; Treatment adherence results not provided (only reference to audiotapes); Most participants exposed to confounds with no control for variables; 1 session; Level of therapist training not reported; Use of a comparison but no control; Low statistical power</td>
</tr>
<tr>
<td>9. Macklin, Metzger, Lasko, Berry, Orr, &amp; Pitman, 2000</td>
<td>EMDR (n = 13 of 17 from original Pitman et al., 1996 study), Non-EMDR (n = 14)</td>
<td>CAPS, IES, M-PTSD, Symptom Check List-90-Revised Global Symptom Inventory</td>
<td>Effect size worsened over the five-year follow up</td>
<td>DSM diagnosis; Reliable, valid, and adequate measures used; multiple outcome measures</td>
<td>Assessor reliability unclear; Non-EMDR condition non-randomized; Treatment adherence unclear; Nonconfounded conditions unclear; 10.5 (eyes fixed/fingers-tapping) sessions; No qualifications for treating clinicians reported; Low statistical power</td>
</tr>
</tbody>
</table>
The Jensen (1992, 1994) study utilized a sample of 25 Vietnam combat veterans diagnosed with PTSD as conceptualized in the DSM-III-R and measured by the SI-PTSD. They were randomly assigned to one of two groups: EMDR ($n = 13$) or Control ($n = 12$). Treatment was apparently replicable and specific but not the standard EMDR protocol given the limited number of sessions (one intake session and two treatments). Treatment adherence is unclear and the Control group was allowed to seek outside treatment during the course of the study. There is some debate (see Rubin, 2003, 2004; Maxfield, Lake, & Hyer, 2004) regarding this study’s EMDR treatment fidelity and whether the EMDR therapists were properly trained (Alto, 2001). Relatedly however, Davidson and Parker (2001) concluded in their meta-analysis that the existing evidence does not suggest that therapists trained by the EMDR Institute produce superior outcome effects than those that are not. Garb and Boyle (2003, pp. 20–21) state:

> ‘In empirical studies, clinicians have rarely been more accurate than graduate students, regardless of the type of information provided to clinicians. Studies have revealed no differences in accuracy between experienced clinicians and graduate students when judgments are made on basis of interview data... biographical and history information... behavioral observation data... data from therapy sessions... and all of the data that clinical and counseling psychologists usually have available in clinical practice’.

Rosen (1999) also notes that proponents of EMDR have not been consistent in their own recommendations regarding the importance of treatment fidelity. Therefore the issue of precise adherence to EMDR treatment protocols in the Jensen (1992, 1994) study may be a moot point. The instruments used to assess outcomes at post-treatment were the M-PTSD and the SI-PTSD. It is unclear whether or not the assessor was blinded or received appropriate training to administer the instruments. On both the SI-PTSD and the M-PTSD post-test, the mean of the EMDR group did not differ significantly from the mean of the control group. The author notes that the ‘subjects in both groups continued to be in significant distress at the end of the study with regard to PTSD symptomatology’ (Jensen, 1994, p. 321). Significant limitations to this study included lack of detail concerning assessments and whether or not the conditions were unfounded, frequency of treatment sessions, and low statistical power.

The Silver et al. (1995) study utilized a sample of 83 Vietnam combat veterans. It is assumed that all veterans received a formal diagnosis of PTSD given the study was conducted as part of a VA inpatient PTSD program evaluation. Clients were assigned (non-randomly) to one of four groups: EMDR ($n = 13$), Relaxation ($n = 9$), Biofeedback ($n = 9$), and Control ($n = 55$). It is unclear what the length of treatment was. It is also unclear what level of training the therapist had received, though it is indicated that he was a doctoral student with 10 years of experience with this
population. Treatment appears to have been replicable, but not standard EMDR protocol. Treatment adherence is not reported. The primary outcome measure was a non-standardized, self-report problem report form (PRF). It was not reported if the assessor was trained to administer the instrument. Non-standardized outcome measures were used and indicated that the combat veterans treated with EMDR did better than the control group clients across all variables measured; EMDR did better than the Relaxation group in all variables with the exception of Depression; and better than Biofeedback group across all variables. However, Rubin (2003) notes that ‘multiple t-test comparisons inflated the Type I error probability (Lohr, Lilienfeld, Tolin, & Herbert, 1999; Lohr, Kleinknecht, Tolin, & Barrett, 1995)’ (p. 7). Significant limitations to this study included the fact that clients were not randomly assigned, the use of an outcome measure of unknown reliability and validity, having a single non-blind therapist providing or supervising all treatments, and a lack of statistical power.

Boudewyns and Hyer (1996) utilized a sample of 61 combat veterans diagnosed with DSM-III-R-defined PTSD (American Psychiatric Association, 1987) and measured by the SCID, the War Stress Inventory (Fontana, Rosenheck, & Spencer, 1990), and the Combat Exposure Scale. They were randomly assigned to one of three groups: EMDR (n = 21), Exposure (n = 18), and Control (n = 22). It is unclear what level of EMDR training the therapists (all licensed psychologists) possessed. Exact treatment fidelity is unknown, though the therapists were videotaped and at least two tapes were reviewed and scored by unnamed EMDR experts. Supervisory feedback was provided to the therapist. Treatment appears to have been replicable and specific but did not follow EMDR protocol (only five to seven sessions). Treatment adherence is unclear. The authors attempted to control for confounds by weaning participants off of psychotropic medications and excluding those ‘with a co-existing Axis I diagnosis of Organic Mental Disorder, Schizophrenia, or with a diagnosis in the category ‘Psychiatric Disorders not Elsewhere Classified’ and an Axis II diagnosis of Antisocial Personality Disorder’ (Boundwyns & Hyer, 1996, p.188). The instruments used to assess post-treatment outcomes were the CAPS and the IES. The assessor was blinded; however, it is unclear whether or not the assessor was trained and supervised in the administration of the instruments. While means were provided for both the CAPS and the IES, the study authors did not include standard deviations, thus limiting the ability to calculate effect sizes. The IES post-test did not show significant change. The CAPS post-test showed a significant drop in all three groups; however, given that there were no significant interactions the authors note that ‘it may be argued that exposure alone is the operative component’ (Boudewyns & Hyer, 1996, p. 190). A significant limitation to the usefulness of this study is the lack of standard deviations provided.

Pitman et al. (1996) utilized a sample of 17 Vietnam combat veterans diagnosed with PTSD as defined by the DSM-III-R and measured by the SCID. They were
randomly assigned to receive either eye-movement or eyes-fixed components of EMDR for the first series of weekly sessions. On the second week they received the alternate component. The therapists included three doctoral-level behavior therapists trained in EMDR at Level II. Exact treatment fidelity was considered to be in the low to moderate acceptable range as determined by a review of videotaped sessions by a purported expert in EMDR. Feedback was provided to the therapist. The treatment was apparently replicable and specific but not consistent with the standard EMDR protocol (mean number of session was 9.7; the veterans were required to finger-tap in a non-bilateral manner instead of bilaterally tracking the therapist’s finger with their eyes (Rubin, 2003). Treatment adherence is unclear, though the authors state that ‘participants in the present study who expressed concern that the therapy was having negative effects were not specifically encouraged to complete it in the hopes of an eventual positive result’ (Pitman et al., 1996, p. 422). The authors attempted to control for potential confounds by excluding participants ‘with an organic, psychotic, manic, or melancholic disorder, or with current alcohol or other substance dependence’ (p. 421) and/or if they were receiving a non-supportive concurrent treatment. No changes in prescribed medication regimens were changed 1 month prior to the conclusion of the study. The outcome instruments employed were the CAPS, the IES, the M-PTSD, and the Symptom Check List-90-Revised Global Symptom Inventory (Derogatis, 1983). The assessor was blinded and received training to administer the instruments. The authors report a modest to moderate 23% average overall improvement across both component groups; however, no data (means, standard deviations, correlations, etc.) were provided to corroborate this finding. The authors also note that ‘data do not support a role for eye movements in emotional processing during EMDR’ (Pitman et al., 1996, p. 426). Significant limitations to this study included unclear treatment adherence, incomplete data presented, and low statistical power exacerbated by attrition in the EMDR condition.

Carlson et al. (1998) utilized a sample of 35 combat veterans diagnosed with DSM-IV- PTSD (American Psychiatric Association, 1994) and determined by reviewing individual medical records along with the CAPS, the M-PTSD, the IES, and the PTSD Symptoms Scale ‘(a 0- to 10-point global scale devised by the authors that provided participants their only opportunity to self rate their overall symptom status)’ (Carlson et al., 1998, p. 6). Clients were randomly assigned to one of three groups: EMDR ($n = 10$), Biofeedback ($n = 13$), or Routine ($n = 12$). The study reports that the therapist was trained, but the level of training is unclear. Treatment fidelity was variable or self-monitored by therapist only. The treatment was apparently replicable and appears to follow the standard EMDR protocol, including the full 12 treatment sessions. Treatment adherence was unclear as is whether or not there were non-confounded conditions. The instruments used to assess post-treatment changes and the 3 month and 9 month blind follow-up outcomes were the CAPS, the IES, the
M-PTSD, and the PTSD Symptoms Scale. The assessor was independent, but not blinded, and it is unclear whether or not adequate training was provided in using the assessment instruments. All of the outcome measures post-test showed significant improvement in the EMDR condition. The authors conclude that ‘EMDR may be an approach to resolving traumatic experiences that is less threatening for some patients with combat PTSD, impacting positively on attrition across the course of treatment, and helping to assure adequate exposure and cognitive restructuring’ (pp. 22–23). Significant limitations to this study included lack of detail regarding assessor reliability, treatment adherence, whether or not conditions were confounded, and low statistical power.

Devilly, Spence, and Rapee (1998) utilized a sample of 51 Vietnam combat veterans diagnosed with DSM-III-R-defined PTSD and measured by the PTSD-I. Stratified randomization was employed assigning the first 20 clients to either EMDR or EMDR treatment minus the eye movements (REDDR), before receiving approval to implement a ‘standard psychiatric support’ control group (SPS). The next 10 consecutive clients were selected to receive the SPS treatment with the following 21 clients being randomly assigned to one of the three conditions. The M-PTSD was used to measure the outcome. The therapist had completed Level II EMDR training. The treatment was apparently replicable and specific but not standard EMDR protocol due to the limited number of sessions (only two). Treatment adherence is unclear as is whether or not there were non-confounded conditions. The instrument used to assess post-treatment outcomes was the M-PTSD. It is unclear whether the assessor was trained and supervised in the administration of the instrument. On the M-PTSD post-test, the authors report that ‘no statistically significant differences were found between the three conditions’ (p. 449). A 6 month follow up, with an attrition of three veterans in both the EMDR and REDDR conditions, resulted in further reduction of apparent treatment effects. Significant limitations to this study included a compromised randomization strategy, lack of detail regarding assessor, therapist, low frequency of treatment sessions, and low statistical power.

Rogers et al., (1999) utilized a sample of 12 Vietnam combat veterans diagnosed with DSM-III-R-defined PTSD and measured by the CAPS. They were randomly assigned to one of two groups: EMDR ($n=6$) and Exposure ($n=6$). Multimodal measures were employed. The level of therapist training was not reported. The treatment was apparently replicable and specific but not consistent with the standard EMDR protocol given the limited number of therapy sessions (only one); this is noteworthy given EMDR proponents’ concern regarding dosage adequacy. Treatment adherence appeared to be strong since the authors report that sessions were audio-taped; however, results of fidelity were not reported. It is unclear whether or not there were non-confounded conditions. The instrument used to assess post-treatment outcomes was the IES and the participants were ‘instructed to complete the inventory
as it related to their most distressing memory’ (Rogers et al., 1999, p. 126). The authors note that it is likely other researchers have not instructed their participants to do this; rather, they focused on their overall war experience. The IES post-test showed improvement in both conditions for the Intrusion subscale and no changes for the Avoidance subscale. Significant limitations include number of treatment sessions and small sample size.

Macklin et al. (2000) study is a 5-year follow-up to the 1996 study conducted by Pitman et al. It utilized a sample of 27 Vietnam combat veterans, with 13 of the original 17 included in the EMDR condition and 14 non-randomly assigned to the non-EMDR condition. The instruments employed to assess post-treatment outcomes were the CAPS, the IES, the M-PTSD, and the Symptom Checklist-90-Revised Global Symptom Inventory. The assessor was blinded; however, it is unclear whether or not training was administered in how to administer the instruments. The authors report effect sizes for the CAPS, the M-PTSD, and the Symptom Check List-90-Revised Global Symptom Inventory indicating that PTSD symptoms worsened; however, they indicated that the effect sizes for IES might represent meaningful treatment effects given sufficient statistical power. Significant limitations included lack of detail concerning assessor, treatment adherence, conditions, and small sample size.

DISCUSSION

This review critically examined and summarized the outcomes, strengths and limitations of six experimental and three quasi-experimental studies of varying methodological quality. Findings from these studies present very limited evidence supporting the effectiveness of EMDR in reducing PTSD in combat veterans. Placebo responses to believable but inert psychosocial treatments can be exceedingly powerful (Kirsch, 2005; Price, Finniss, & Benedetti, 2008), and EDMR bears many of the characteristics of being, at least in part, a placebo treatment overlaid on standard behavioral practices (Rosen, Lohr, McNally & Herbert, 1998). There are no well-designed RCTs comparing EMDR against real-life exposure therapy, a treatment with a much stronger level of empirical support in the treatment of PTSD, or, for that matter, against credible placebo-controlled therapies. Given the incidence of combat-related PTSD, the costs it exacts upon the individual experiencing the disorder, their families, and the military, we believe that such well-designed intervention research is urgently needed. Parsimony suggests that EMDR ‘...is simply a version of exposure therapy, and that the eye movement procedure is not essential’ (Najavits, 2007, p. 517), and we are inclined to support this view, at present.

While the authors of included studies most likely did the best with the resources available, all of the studies are limited by small sample sizes and a lack of precision,
increasing the likelihood that random error exerts a strong influence in these studies and decreases their sensitivity to detect legitimate treatment effects. Higher power also helps reduce Type II errors. The published studies are further limited by a lack of detail and transparency in reporting data. Future work in this area should include both means and standard deviations (or correlations, odds-ratios, risk ratios, etc.) for the outcome measure(s) so that an effect size can be calculated. Better description of participant and therapist information is also suggested to allow the reader to better understand potential (or lack of) biases and attrition rates. Few follow-up assessments were conducted over long time periods, raising the possibility that some initial improvements may have been brief, of that some initial lack of improvements were followed by subsequent changes.

Another limitation in studying response to psychotherapy in clients with PTSD stemming from exposure to military combat is the finding that client self-reports of such combat exposure are often not able to be corroborated by examining their military records (Frueh et al., 2005), suggesting that combat exposure may be inaccurate and over-reported by the clients themselves. We suggest that future studies of the effectiveness of various psychotherapies on presumptively combat-related PTSD be limited to military veterans whose service records adequately document their exposure to combat.

We specifically note that the available evidence regarding the efficacy of EMDR as a treatment for PTSD occurring in military combat veterans does not meet the threshold required to be labeled as an empirically supported treatment, as defined by the criteria established by Division 12, Section III of the American Psychological Association. These standards require, among other features:

‘At least two good between-group design experiments demonstrating efficacy in one or more of the following ways: A. Superior (statistically significant) to pill or psychological placebo or to another treatment. B. Equivalent to an already established treatment in experiments with adequate sample size’ (see Chambless et al., 1998, 3–16).

This lack of evidence raises concerns about the apparent premature adoption of EMDR as an officially endorsed treatment for members of the military and veterans diagnosed with PTSD by the Department of Defense and the Department of Veteran Affairs (see Russell, Silver, Rogers & Darnell, 2007). History tells us that some forms of psychosocial treatment can have iatrogenic or nociceptive effects, deleterious consequences quite separate from the possibility that they are actually inert or ineffective. To the extent that the victims of combat-related PTSD deserve to receive therapies consistent with evidence-based practice, it can be currently seen as premature to offer EMDR except as an experimental treatment provided in the context of well-designed and properly controlled outcome studies.
While the PS scale was utilized to review these articles critiqued in this review, no standard scale development techniques (Devellis, 2003; Spector, 1991) were presented by Hertlein and Ricci (2004). This suggests that there is a need for either adapting a tested scale, such as the 22-item checklist developed through the consolidated standards of reporting trials (CONSORT) initiative (Greenfield et al., 2005) or creating and testing one designed specifically for EMDR studies. Regardless, the need for an evaluative tool for readers to assess the validity, limitations, and potential generalizability of randomized controlled trials of psychosocial treatments is needed.

One concluding point – the studies we reviewed involved samples of combat veterans predominantly from the Vietnam war. However, the current combat veteran cohort is unlike any other in part due to multiple deployments, longer rotations in theater, a higher number of females exposed to combat situations, and greater numbers of National Guard and Reserve units. It is worth further research to determine if EMDR or other cognitive-behavioral therapies are effective with this cohort of combat veterans.

REFERENCES

*Refers to included study.


Copyright © 2009 John Wiley & Sons, Ltd.

*Behav. Intervent. 25*: 1–19 (2010)

DOI: 10.1002/bin


