Assessing the Long-Term Effects of EMDR: Results from an 18-Month Follow-Up Study with Adult Female Survivors of CSA

Tonya Edmond
Allen Rubin

ABSTRACT. This 18-month follow-up study builds on the findings of a randomized experimental evaluation that found qualified support for
the short-term effectiveness of Eye Movement Desensitization and Reprocessing (EMDR) in reducing trauma symptoms among adult female survivors of childhood sexual abuse (CSA). The current study provides preliminary evidence that the therapeutic benefits of EMDR for adult female survivors of CSA can be maintained over an 18-month period. Furthermore, there is some support for the suggestion that EMDR did so more efficiently and provided a greater sense of trauma resolution than did routine individual therapy.

KEYWORDS. EMDR (Eye Movement Desensitization and Reprocessing), childhood sexual abuse, treatment effectiveness, trauma

Francine Shapiro (1989) developed Eye Movement Desensitization and Reprocessing (EMDR) primarily as a treatment for trauma. It is a treatment approach aimed at assisting individuals to be able to tolerate traumatic memories and rationally process traumatic information in a productive manner. EMDR consists of eight phases and is accompanied by specific protocols for different treatment issues. The eight phases of EMDR proposed by Shapiro (1989) are as follows: (1) gathering information about the client’s history and developing a solid treatment plan; (2) preparing the client for the work to be done; (3) assessing all components of the treatment target; (4) desensitizing the targeted traumatic material through the use of eye movements or an alternate type of stimulation; (5) installing the identified positive cognition; (6) scanning the body to locate any unresolved residual material; (7) closure; and (8) reevaluation.

EMDR is based on the notion that within the context of the treatment approach, eye movements, as well as some alternative forms of bilateral stimulation (e.g., finger snaps near the ears, right and left hand taps, and shoulder taps), enable traumatized individuals to rationally process trauma-related distress and overcome posttraumatic symptoms (Shapiro, 2001). The theoretical framework that Shapiro (2001) developed to explain the observed effects of EMDR has been termed Adaptive Information Processing. She hypothesizes “that the procedural elements of EMDR, including the dual attention stimuli, trigger a physiological state that facilitates information processing” (p. 31). For a more detailed re-
Shapiro (1989) conducted an initial study on the efficacy of EMDR in a randomized experiment on 22 subjects who were experiencing traumatic memories connected to either the Vietnam War, childhood sexual or emotional abuse, or a sexual or physical assault. Those who received brief EMDR therapy reportedly experienced a significant reduction in reactions to their traumatic memories, as well as dramatic improvements in their cognitions, whereas no such improvements were noted for the control group. One common criticism of Shapiro’s (1989) efficacy study, however, involves her failure to utilize standardized outcome measures, which left the results of this initial study vulnerable to the influence of demand characteristics and experimenter expectancy effects (Acierno, Hersen, Van Hasselt, Tremont, & Meuser, 1994; Greenwald, 1994; Herbert & Mueser, 1992; Lohr, Kleinknecht, Tolin, & Barrett, 1995). Since this initial attempt to demonstrate the efficacy of EMDR, a number of controlled studies have been conducted in response to the early criticisms. These more controlled studies have also found evidence suggesting the efficacy of EMDR. Even among EMDR’s most ardent critics, the dramatic increase in methodological rigor employed by the more recent studies investigating EMDR has been recognized (Lohr, Tolin, & Lilienfeld, 1998).

Most of the research to date supporting the effectiveness of EMDR has involved its application to a single traumatic memory, or a single memory cluster (i.e., a single traumatic memory and its associated memories) and has focused on the reduction or elimination of symptoms related to Posttraumatic Stress Disorder (PTSD) (Jensen, 1994; Marcus, Marquis, & Sakai, 1997; Renfry & Spates, 1994; Rothbaum, 1997; Shapiro, 1989; Wilson, Becker, & Tinker, 1995). In addition, Van Etten and Taylor (1998) conducted a meta-analysis of 59 PTSD treatment outcome trials and concluded that EMDR was equal in effectiveness to behavior therapies for treating PTSD symptoms and both approaches were found to be superior to other psychotherapies or pharmacotherapies in effectively treating this disorder.

Further support of EMDR has come from the endorsements it has received from professional associations striving to develop guidelines for the treatment of trauma. Chambless and colleagues (1998), for example, developed a list of empirically validated therapies that they categorized as either well-established treatments or probably efficacious treatments. They categorized EMDR as probably efficacious for treating civilian PTSD. Likewise, the International Society for Traumatic Stress Studies
has identified EMDR as an effective treatment for PTSD in their current treatment guidelines (Chemtob, Tolin, van der Kolk, & Pitman, 2000).

Despite the support that has been found for this treatment approach, debate about the efficacy of EMDR continues. In two different reviews of the treatment literature for PTSD, for example, both Keane (1998) and Foa and Meadows (1997) reported mixed results concerning its efficacy and methodological limitations, which have prevented them from endorsing EMDR as an effective treatment of PTSD. Rather, both posit a need for additional well-designed, controlled studies before any determination about the efficacy of EMDR could be made. In addition, Lohr et al. (1998), in an impressive critique of 17 studies, concluded; “the process of empirical validation has revealed a disparity between the data justifying the application of EMDR and its current widespread use” (p. 149). These authors suggest that the effects purportedly attributed to EMDR are more likely the product of exposure and/or nonspecific treatment effects. Despite such criticism, however, the usage of EMDR continues to grow. Presently, close to 45,000 psychotherapists around the world have received training in this treatment approach (R. Dunton, personal communication, June 13, 2003).

Although most of the research on EMDR has involved traumatized target groups, only one controlled experiment thus far has evaluated its effectiveness in reducing trauma symptoms exclusively with adult survivors of childhood sexual abuse (CSA). This study, conducted by Edmond, Rubin, and Wambach (1999), found a discrepancy between the short-term and long-term effects of EMDR treatment. Edmond et al. (1999) randomly assigned 59 female CSA survivors to one of three treatment conditions: (a) individual EMDR treatment; (b) routine individual treatment; or (c) a delayed treatment control group. Those placed in one of the two treatment conditions received six weekly 90-minute individual treatment sessions focused on the CSA issue seen as most troubling by the participant. According to Edmond et al. (1999), therapy in the routine treatment condition involved a variety of methods, techniques and theories that were incorporated into an approach best suited to address the therapeutic target introduced by each participant. For example, the therapists in the study utilized aspects of cognitive-behavioral and psychodynamic theories, employing techniques like cognitive restructuring, behavior modification, relaxation, ego strengthening, interpretation and dream work. This “eclectic” type of approach was selected because survivors do not generally exhibit a single specific clinical symptom, but rather tend to present with multiple symptoms and dysfunctional behaviors. Thus, most treatment approaches used in practice are multistage and multimodal (Gordon & Alexander, 1993),
which allows practitioners to address a broad range of traumatic targets presented by adult survivors of childhood sexual abuse.

In Edmond et al.’s 1999 study, participant functioning was measured on four standardized instruments at pretest, posttest (6 weeks), and follow-up (3 months). The four instruments comprising the outcome measure included: (a) The State-Trait Anxiety Inventory (STAI); (b) The Impact of Events Scale (IES); (c) The Beck Depression Inventory (BDI); and (d) The Belief Inventory (BI). Results of this study indicated that both EMDR and routine individual therapy were effective in reducing the targeted trauma symptoms compared to the control condition (i.e., delayed treatment). At posttest, EMDR was found to be equally effective in reducing the target symptoms compared to routine individual therapy for this population. However, at follow-up, the EMDR group displayed significantly better outcomes compared to those in the routine individual treatment group on two of the four standardized measures, trauma-specific anxiety and depression. While the EMDR group demonstrated no clinically significant symptoms of anxiety or depression, the routine individual therapy group did. These results suggested that EMDR might be producing more enduring trauma resolution than routine individual therapy, which prompted the need to conduct a longer follow-up study.

Consequently, the present study represents an attempt to meet this need by evaluating the progress of the CSA survivors studied by Edmond et al. (1999) in a follow-up assessment conducted eighteen months after the completion of treatment. This strategy allowed the present study not only to monitor participants’ progress, but also enabled another comparison of EMDR’s short- and long-term effects to be made. It was hypothesized that the therapeutic gains demonstrated by those in the EMDR treatment condition at 3 months would be maintained at 18 months as well. In addition, it was further hypothesized that those receiving EMDR would continue to reflect a better resolution of their trauma symptoms compared to those who received individual routine therapy.

**METHODS**

**Participants**

Of the 59 female survivors that were included in the original study conducted by Edmond et al. (1999), 42 (71%) were included as participants in the present study. Eighty-three percent of the participants in the
18-month follow-up study (N = 35) identified themselves as Caucasian, 5% (N = 2) as African American, 2% (N = 1) as Hispanic, 2% (N = 1) as Asian American, 2% (N = 1) as mixed ethnicity, and 5% (N = 2) identified themselves as “other.” Seventy-seven percent (N = 7) of the non-Caucasian participants who participated in the original study conducted by Edmond et al. (1999) also participated in the present follow-up study, suggesting that ethnicity was not a factor in treatment motivation, compliance, or dropout. The mean age of participants was found to be 36 years (range = 18-51), and participants possessed a mean of 15.2 years of education. Sixty-two percent (N = 26) of the participants were employed full-time, 14% (N = 6) were employed part-time, and their mean income was $30,400. Thirty-five percent (N = 15) of the women were married, 23% (N = 10) divorced, 23% (N = 10) single, 14% (N = 6) living with a significant other, 5% (N = 2) were widowed, and 45% (N = 19) had children. The vast majority of participants, 91% (N = 38) had obtained some type of therapy to address the sexual abuse prior to participating in the study.

The CSA survivors in this study reported severe histories of abuse. The mean age at which the abuse began was 6, and the mean age in which it ended was 13. Nearly half of the survivors, 48% (N = 20), endured the abuse for 5 or more years. The abuse occurred between three and four times a month to three to five times a week for 61% (N = 26) of participants. Ninety-eight percent (N = 41) of participants identified their perpetrators as family members, with males most often identified as the perpetrators—biological fathers (49%; N = 21), brothers (21%; N = 9), grandfathers (17%; N = 7), and stepfathers (12%; N = 5). During the course of the abuse, 44% (N = 18) of participants reported being sexually abused by more than one perpetrator. In addition, 64% (N = 27) were physically abused within their family of origin, and 64% (N = 27) experienced some form of revictimization in adulthood such as sexual assault or domestic violence.

Seventeen (40%) of the participants in the present study originally belonged to the routine treatment group, 14 (33%) to the EMDR group and 11 (26%) to the control group. In total, 62% (N = 26) of the eighteen month follow-up sample obtained additional therapy after post-testing: 65% (N = 11) originally assigned to the routine treatment group, 57% (N = 8) of those originally assigned to the EMDR group, and 64% (N = 7) of those from the original control group. The types of additional treatment received included individual, couple’s, and group therapy. Within the routine treatment group, 65% (N = 11) of participants obtained individual therapy, 18% (N = 2) obtained couple’s therapy, and 27% (N = 3)
participated in support groups. Of these 11 participants who received additional individual therapy, 45% \((N = 5)\) received EMDR. Of the participants who were originally assigned to the EMDR group, 43\% \((N = 6)\) obtained additional individual therapy, 7\% \((N = 1)\) attended couple’s therapy, and 14\% \((N = 2)\) participated in groups. In terms of the participants originally assigned to the control group, 36\% \((N = 4)\) obtained individual therapy, 75\% \((N = 3)\) of which received EMDR during the course of their treatment, and 27\% \((N = 3)\) attended couple’s therapy. Given the subsequent treatment obtained, the original control group represented an additional comparison treatment group in the present study.

**Measures**

The State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). The state anxiety scale of this self-report measure was used to assess anxiety related to any trauma-specific issue of concern. Test-retest reliability is relatively low for the S-Anxiety scale, which is as it should be for assessing situational stress. The internal consistency for the scale is very high, with median alpha coefficients of .90. The STAI has good construct validity as demonstrated by its ability to discriminate between normal and psychiatric patients with anxiety symptoms (Spielberger et al., 1983).

The Impact of Events Scale (IES; Horowitz, Wilner, & Alvarez, 1979). This self-report measure was used to assess posttraumatic stress symptoms for any specific trauma. The IES has known groups’ validity and very good internal consistency, with alphas ranging from .79 to .92.

The Beck Depression Inventory (BDI; Beck & Steer, 1993). This commonly used self-report measure was utilized by the present study for the purpose of assessing clinically significant levels of depression ranging from mild to severe. The BDI has been in use for 30 years and has been shown to have good-to-excellent validity and reliability. Strong internal consistency has been demonstrated with split-half reliabilities, which have ranged from .78 to .93. Test-retest reliability has been good, with a range of .48 for psychiatric patients after three weeks to .74 for college students after four months. The BDI has also been shown to have good-to-excellent concurrent validity, as demonstrated by significant correlations with other depression measures as well as with clinician’s ratings of depression (Corcoran & Fischer, 1987).
The Belief Inventory (BI; Jehu, Gazan, & Klassen, 1985). This self-report measure identifies and measures common distorted beliefs among adult survivors of childhood sexual abuse. The inventory has very high test-retest reliability (.93) after one week. It has face validity as well as concurrent validity that has been established with the Beck Depression Inventory (.55) (Jehu, Gazan, & Klassen, 1985).

Design and Procedures

Participants were obtained from the sample studied in Edmond et al.’s original 1999 study, which involved an experimental design with random assignment to either EMDR, Routine Individual Therapy or a delayed treatment control group. Participants of this original study were recruited through the use of a newspaper ads and flyers provided to various practitioners located in central Texas. To be selected for the study, respondents had to be adult female survivors of childhood sexual abuse who had specific memories of their abuse, had no previous exposure to EMDR, exhibited no contraindications for use of EMDR (that is, ocular problems, active suicidal ideation, serious medical condition, inadequate ego strength, or severe mental disorders such as psychosis), and were not receiving any concurrent therapy.

The principal investigator, who had conducted the pre- and post-testing in Edmond et al.’s (1999) original study, was the only person that conducted the testing for the eighteen-month follow-up analyzed in the present study. Each participant that could be located was contacted by phone and each agreed to participate in an 18-month follow-up interview, which was conducted in person. As the funding for this study was limited, no compensation could be offered to the participants. In addition to the standardized measures that the Principal Investigator administered, she also obtained measures of the Subjective Units of Discomfort Scale (SUDS) and the Validity of Cognition Scale (VOCs). These subjective instruments were selected because they are part of the standard administration of EMDR and have been used in numerous EMDR studies as the primary method of measuring reported change. The SUDS (Wolpe, 1990) is used to obtain a verbal report from subjects about their level of emotional distress associated with a traumatic experience. The VOC (Shapiro, 1989) is used in EMDR to rapidly assess the client’s cognitive beliefs associated with the trauma. Therapists in the original Edmund et al. (1999) study also obtained these measures as in-session process measures, thereby allowing the authors of the present study to compare participants’ subjective sense of improvement at 18 months with those obtained earlier as well.
RESULTS

A chi-square test found no relationship between ethnicity and whether members of Edmond et al.’s (1999) original sample agreed to participate in the present follow-up study \( (p = .70; V = .05) \). Likewise, no such relationship was found on the basis of participants’ income, marital status or number of children. However, the subset of survivors who were included as participants in the present study were found to differ from the subset of participants in Edmond et al.’s (1999) original sample that did not participate in the present study on two demographic characteristics: age and education level. Those who participated in the present study were found to have a mean age of 36 years, while those who did not participate were found to have a mean age of 30 years \( (p < .05) \). Also, survivors who participated in the present study were found to possess a mean of 15.2 years of education, compared to the mean 13.5 years of education found for those who did not participate in the present study \( (p < .01) \).

Participants in the present study were not found to significantly differ in terms of whether additional therapy was obtained or in the type of therapy (individual, couples, or group) received on the basis of the treatment condition they were originally assigned to (i.e., routine treatment, EMDR, or control). The EMDR group had a mean of six sessions of therapy between post-testing and the 18-month follow-up, with a range of 0 to 24 sessions. The routine treatment group had a mean of 20, with a range of 0 to 64 sessions between post-testing and the eighteen-month follow-up. Those in the control group attended an average of 13 sessions between post-testing and the 18-month follow-up, with a range of 0 to 72. Although the control group obtained an average of twice as many therapy sessions as the EMDR group and the routine treatment group obtained an average of three times as many therapy sessions as the EMDR group had, this difference did not reach statistical significance \( (p = .065) \).

More than half \( (N = 14) \) of the total number of participants who had received subsequent therapy \( (N = 26) \) reported that they focused on the same target issue that had been addressed in the original study conducted by Edmond et al. (1999), indicating that the target issue had not been adequately resolved within the six sessions of therapy provided during the original study. This appears particularly true for the routine treatment participants, nine of whom focused on the same target issue in their subsequent therapy sessions, whereas only one participant originally assigned to the EMDR condition and four participants who had

Tonya Edmond and Allen Rubin 77
originally been controls continued to focus on the same treatment issue. This difference proved to be statistically significant \((n = 26; p = .01)\).

**Maintenance of Therapeutic Gains**

The first hypothesis of the present study suggested that the therapeutic gains demonstrated by those in the EMDR treatment condition at 3 months would be maintained at 18 months. Of the 52 participants whose data was analyzed by Edmond et al. (1999), 42 were included as participants in the present study and were, therefore, tested in an 18-month follow-up assessment. However, the data of 62\% \((N = 26)\) of these participants was eliminated from this analysis due to the additional therapy that was received by this subset of participants following the termination of Edmond et al.’s (1999) study. The influence of this additional therapy made accurately assessing the enduring effects of Edmond et al.’s (1999) original treatment conditions difficult for these participants. Consequently, the investigation of the first hypothesis was based on an analysis of the data provided by six participants who had originally been placed in the routine treatment condition, six participants who had originally been placed in the EMDR condition, and four of the original controls, none of whom obtained subsequent treatment following completion of the posttest. Given the extremely small sample size \((N = 16)\), paired-sample t-tests were conducted to examine within group stability of scores from posttest to the 18-month follow-up. Separate t-tests were run for each of the standardized trauma symptom measures used in the study, as well as for the SUDS and VOC.

The t-test results indicated that none of the mean scores on the standardized trauma symptom measures (BDI, STAI, IES, BI) were significantly different from the posttest to the 18-month follow-up for any of the treatment conditions (see Table 1). With only 16 subjects, the statistical power for these t-tests was less than .25 (Rubin & Babbie, 2001). However, the data in Table 1 obviate this concern, because the EMDR group improved slightly on every standardized measure between post-testing and 18-month follow-up. Since the purpose of the t-tests was to see if therapeutic gains were maintained, a Type 2 error would be of concern primarily if, instead of improving slightly, the scores deteriorated between post-testing and follow-up. Therefore, it can be concluded that both the EMDR and routine treatment groups appeared to maintain the therapeutic gains that had been observed at posttest 18 months after the original six sessions of therapy were provided. In contrast, the scores of the control group deteriorated somewhat between
posttest and follow-up on three of the four standardized measures (see Table 1). In addition, scores deteriorated somewhat for the EMDR and routine treatment groups on each of the two subjective measures, but only for the routine treatment group’s VOC scores was this difference statistically significant ($t = 2.739; p = .04$).

**Standardized Outcome Measures**

In order to investigate the hypothesis that those receiving EMDR would continue to reflect better resolution of their trauma symptoms compared to those who received individual routine therapy, a Multivariate Analysis of Covariance (MANCOVA) was employed to test the overall significance of the differences in eighteen-month follow-up scores among the three treatment groups across all four standardized outcome measures (STAI, IES, BDI, and BI). Subsequent therapy, type of therapy, and number of sessions obtained between post-testing and the eighteen-month follow-up were each used individually as covariates, however, and none proved to be statistically significant. Consequently, all 42 women who participated in the 18-month follow-up were included in this analysis. Differences in 18-month follow-up scores between treatment conditions were found to be significant ($N = 42$; Wilks’ Lambda = .652; $F = 2.084; p = .049$). Given the significant differences between groups demonstrated through MANCOVA, a separate univariate analysis of variance was conducted for each measure. Significant differences were found for all four dependent variables at the .05 level.

In Edmond et al.’s (1999) original study, those assigned to the EMDR treatment condition had lower scores (fewer clinical symptoms) on all four of the outcome measures at the three-month follow-up compared to those in the routine treatment condition. Those differences were statistically significant ($p < .05$) for two of the four standardized measures—trauma-specific anxiety and depression. At the 18-month follow-up, while the EMDR group continued to produce lower scores (i.e., endorsed fewer symptoms) on all four of the dependent variables than the routine treatment group, no significant univariate differences between the treatment groups were found. The significant differences found on the univariates were between the EMDR group and the control group. On all four dependent measures, the EMDR group had significantly better scores than the control group, while the routine treatment group was not significantly better than either EMDR or the control group on any of the dependent variables (see Table 2).
A MANCOVA was also used to test the overall significance of the differences in eighteen-month follow-up scores between the EMDR group and the routine treatment group on the subjective process measures (i.e., the SUDS and VOC). Only the EMDR and routine treatment groups (N = 31) have SUDS and VOC scores, since they were process measures used within both treatments, and therefore, the control group could not be included in this analysis. The differences between the two treatment groups were not found to be significant for either of the subjective measures.

**Subjective Outcome Measures**

**TABLE 1. Mean Scores, Standard Deviations, and Sample Size on All Measures, by Group and Time for Participants Who Did Not Obtain Subsequent Therapy After Posttest**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group</th>
<th>Pretest***</th>
<th>N</th>
<th>Posttest***</th>
<th>N</th>
<th>3-Month*** Follow-Up</th>
<th>N</th>
<th>18-Month Follow-Up</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean Score (SD)</td>
<td></td>
<td>Mean Score (SD)</td>
<td></td>
<td>Mean Score (SD)</td>
<td></td>
<td>Mean Score (SD)</td>
<td></td>
</tr>
<tr>
<td>BDI</td>
<td>EMDR</td>
<td>16.3 (7.2)</td>
<td>6</td>
<td>8.6 (6.7)</td>
<td>6</td>
<td>9.2 (10.3)</td>
<td>6</td>
<td>7.2 (10.4)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>15.0 (7.0)</td>
<td>6</td>
<td>12.0 (11.3)</td>
<td>6</td>
<td>11.0 (7.8)</td>
<td>6</td>
<td>11.0 (12.7)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>19.8 (3.6)</td>
<td>4</td>
<td>17.3 (9.8)</td>
<td>4</td>
<td>15.7 (16.5)</td>
<td>4</td>
<td>17.6 (10.8)</td>
<td>4</td>
</tr>
<tr>
<td>STAI</td>
<td>EMDR</td>
<td>58.0 (5.5)</td>
<td>6</td>
<td>31.8 (11.4)</td>
<td>6</td>
<td>33.5 (8.8)</td>
<td>6</td>
<td>31.7 (13.6)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>58.7 (9.4)</td>
<td>6</td>
<td>36.7 (15.1)</td>
<td>6</td>
<td>43.0 (14.7)</td>
<td>6</td>
<td>38.7 (16.1)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>58.8 (13.6)</td>
<td>4</td>
<td>44.8 (9.5)</td>
<td>4</td>
<td>43.3 (23.5)</td>
<td>4</td>
<td>54.0 (16.7)</td>
<td>4</td>
</tr>
<tr>
<td>IES</td>
<td>EMDR</td>
<td>45.3 (15.9)</td>
<td>6</td>
<td>12.0 (15.9)</td>
<td>6</td>
<td>11.7 (15.6)</td>
<td>6</td>
<td>10.2 (13.1)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>26.7 (16.9)</td>
<td>6</td>
<td>14.0 (13.9)</td>
<td>6</td>
<td>27.3 (16.5)</td>
<td>6</td>
<td>16.2 (18.1)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>39.8 (9.1)</td>
<td>4</td>
<td>32.5 (13.4)</td>
<td>4</td>
<td>23.0 (19.3)</td>
<td>4</td>
<td>34.3 (23.5)</td>
<td>4</td>
</tr>
<tr>
<td>BI</td>
<td>EMDR</td>
<td>23.8 (12.2)</td>
<td>6</td>
<td>10.3 (11.0)</td>
<td>6</td>
<td>9.2 (11.6)</td>
<td>6</td>
<td>9.0 (12.4)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>22.2 (11.0)</td>
<td>6</td>
<td>15.7 (9.6)</td>
<td>6</td>
<td>12.8 (7.8)</td>
<td>6</td>
<td>11.0 (9.3)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>41.0 (19.4)</td>
<td>4</td>
<td>35.8 (18.3)</td>
<td>4</td>
<td>19.7 (8.2)</td>
<td>4</td>
<td>22.3 (10.5)</td>
<td>4</td>
</tr>
<tr>
<td>SUDS</td>
<td>EMDR</td>
<td>7.0 (2.4)</td>
<td>6</td>
<td>1.0 (1.3)</td>
<td>6</td>
<td>1.1 (1.0)</td>
<td>6</td>
<td>1.8 (2.1)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>7.7 (2.0)</td>
<td>6</td>
<td>4.3 (3.0)</td>
<td>6</td>
<td>4.5 (2.1)</td>
<td>6</td>
<td>4.2 (3.6)</td>
<td>6</td>
</tr>
<tr>
<td>VOC</td>
<td>EMDR</td>
<td>2.7 (1.4)</td>
<td>6</td>
<td>7.0 (0.0)</td>
<td>6</td>
<td>6.5 (.84)</td>
<td>6</td>
<td>6.0 (.89)</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>2.5 (1.4)</td>
<td>6</td>
<td>5.1 (1.7)</td>
<td>6</td>
<td>5.3 (1.4)</td>
<td>6</td>
<td>4.7 (1.9)</td>
<td>6</td>
</tr>
</tbody>
</table>

***Data from pretest, posttest, and 3-month follow-up assessment was obtained from Edmond et al. (1999)
**TABLE 2. Mean Scores, Standard Deviations, and Sample Size on Standardized Measures, by Group and Time**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Group</th>
<th>Mean Score (SD)</th>
<th>N</th>
<th>Mean Score (SD)</th>
<th>N</th>
<th>Mean Score (SD)</th>
<th>N</th>
<th>Mean Score (SD)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EMDR</td>
<td>16.0 (6.3)</td>
<td>20</td>
<td>10.3* (7.2)</td>
<td>20</td>
<td>4.3* (4.5)</td>
<td>18</td>
<td>5.2* (7.1)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>17.7 (8.9)</td>
<td>20</td>
<td>12.7 (8.2)</td>
<td>20</td>
<td>11.9 (7.1)</td>
<td>16</td>
<td>10.8 (8.9)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>19.3 (7.5)</td>
<td>19</td>
<td>16.7 (10.0)</td>
<td>19</td>
<td>Not included</td>
<td></td>
<td>14.6 (8.5)</td>
<td>11</td>
</tr>
<tr>
<td>BDI</td>
<td>EMDR</td>
<td>58.4 (9.6)</td>
<td>20</td>
<td>34.7* (10.7)</td>
<td>20</td>
<td>30.1* (8.2)</td>
<td>18</td>
<td>30.8* (9.1)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>59.8 (11.4)</td>
<td>20</td>
<td>40.4* (12.2)</td>
<td>20</td>
<td>41.8 (14.4)</td>
<td>16</td>
<td>37.2 (11.0)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>59.4 (12.9)</td>
<td>19</td>
<td>54.0 (17.3)</td>
<td>19</td>
<td>Not included</td>
<td></td>
<td>46.6 (12.8)</td>
<td>11</td>
</tr>
<tr>
<td>STAI</td>
<td>EMDR</td>
<td>38.7 (16.4)</td>
<td>20</td>
<td>14.1* (15.9)</td>
<td>20</td>
<td>10.3 (12.4)</td>
<td>18</td>
<td>8.6* (9.9)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>34.8 (14.6)</td>
<td>20</td>
<td>14.0* (12.0)</td>
<td>20</td>
<td>18.0 (15.1)</td>
<td>16</td>
<td>13.2 (15.9)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>39.6 (12.5)</td>
<td>19</td>
<td>32.1 (17.0)</td>
<td>19</td>
<td>Not included</td>
<td></td>
<td>25.5 (24.5)</td>
<td>11</td>
</tr>
<tr>
<td>IES</td>
<td>EMDR</td>
<td>24.1 (12.2)</td>
<td>20</td>
<td>12.1* (10.0)</td>
<td>20</td>
<td>8.6 (9.0)</td>
<td>18</td>
<td>6.6* (8.2)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Routine</td>
<td>25.4 (14.0)</td>
<td>20</td>
<td>16.3* (10.4)</td>
<td>20</td>
<td>15.3 (12.4)</td>
<td>16</td>
<td>11.8 (10.9)</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>27.1 (15.2)</td>
<td>19</td>
<td>25.1 (15.3)</td>
<td>19</td>
<td>Not included</td>
<td></td>
<td>18.7 (11.2)</td>
<td>11</td>
</tr>
</tbody>
</table>

*aSignificantly better than control at p ≤ .05.
*bSignificantly better than control p ≤ .01.
*cSignificantly better than control at p ≤ .001.
*dSignificantly better than routine individual treatment at p ≤ .01.
*eSignificantly better than routine individual treatment at p ≤ 0.

**DISCUSSION**

In examining the results of this study there is support for the first hypothesis, which stated that the therapeutic gains demonstrated by those who received EMDR treatment in Edmond et al.’s (1999) original study would be maintained 18 months posttreatment. The EMDR group was found to have not only maintained their therapeutic gains, but also to have actually improved slightly on every standardized measure, while the control group’s scores on the outcome measures deteriorated somewhat. Although none of these changes were statistically significant, this trend was still seen as evidence of EMDR’s ability to produce long-term improvements in trauma symptomology. Clinical support, but not statistical support, was also found for the present study’s second hypothesis, which stated that EMDR would continue to reflect better resolution of trauma symptoms compared to those who received routine individual therapy.
Moreover, while both treatment groups appeared to maintain their gains over an 18-month time period, those who received EMDR appear to have done so more efficiently and possibly with a greater sense of trauma resolution. For example, a smaller percentage of the EMDR treatment group obtained additional therapy compared to those who had been placed in the routine therapy and control groups. In addition, those within the EMDR group who did obtain additional therapy received, on average, fewer sessions than those within the routine or control groups. Furthermore, only one survivor within the EMDR group felt the need to focus on the same target issue in the subsequent therapy, compared to nine within the routine group and four within the control group. In fact, even after obtaining additional treatment, both the routine treatment and control group members continued to display clinically significant trauma symptoms that ranged from mild to moderate levels of anxiety, depression, and PTSD. In contrast, the EMDR group had 18-month follow-up scores on the STAI (anxiety), BDI (depression), and IES (PTSD symptoms) that fell in the non-clinical range, reflecting a lack of trauma symptoms. Thus, while participants in both the EMDR and the routine treatment groups may have been able to maintain their therapeutic gains, EMDR treatment appears to have facilitated more resolution.

It was surprising to see that subsequent therapy (an average of 22 sessions) was not a significant covariate in the multivariate analysis of treatment effects on the dependent variables. Given that 46% (N = 19) of the survivors who obtained additional therapy did not focus on the same issue targeted in the original study, it is possible that the subsequent therapy produced effects that the dependent variables were not meant or able to capture. For example, several survivors obtained couples’ therapy. The positive gains associated with such a treatment approach would be more likely to be detected on a measure related to marital problems than on the trauma measures used in the present study. Unfortunately, since the only data obtained by the present study on the subsequent therapy of the participants involved its type and frequency, as well as whether its focus was the same as the original study treatment target, future research may be needed to address this possibility. In addition, the subsequent treatments may not have had an ability to generalize effects to the trauma symptoms that were targeted in the study. In other words, it is possible that whatever was targeted by their subsequent treatment did not spill over and positively effect the symptoms targeted by the present study.

It is also possible that there were other traumatic experiences that generated measurable effects, but that had not yet been adequately ad-
dressed in subsequent therapy. As was done in this study, survivors and their mental health providers often partialize and prioritize treatment to address the issues viewed as most pressing or urgent, which may mean that relevant issues and symptoms are unattended for a period of time. Also, given the abundance of treatments being used that lack empirical validation, it is possible that the additional therapy received was not particularly effective.

It should also be mentioned that long-term follow-up studies on intervention research are few and far between. Furthermore, when follow-up data has been reported in the EMDR literature, as well as with other interventions, very little information has been provided on the potentially confounding interventions received between testing times. Therefore, the present study’s commitment to reporting the intervening interventions obtained and diligence in pointing out the ethical and methodological challenges faced in attempting to assess long-term effects as the result of these intervening interventions represents a unique contribution to the existing literature.

Several methodological limitations of the present study provide grounds for caution in interpreting the findings. In the original study we were only able to obtain 20 clients in each group, which created a big risk for committing a Type II error (i.e., erroneously concluding that the intervention had no effect when in fact it did). The sample size was even smaller in the present follow-up study, which severely limited the power of our multivariate statistics. This limitation might account for the lack of statistically significant findings in the MANCOVA analysis despite the presence of clinically significant differences in the trauma symptoms demonstrated by the two treatment groups.

Furthermore, the possibility of reactive testing effects cannot be ruled out. Limited resources in the unfunded original study influenced the decision to restrict the treatment regimen to six sessions. Longer-term treatment may have been needed to address additional issues. In addition, the possibility that nonspecific treatment factors or placebo effects influenced the results cannot be ruled out. Additionally, the current findings may have been skewed as a result of the unique characteristics of the survivors who participated in the eighteen-month follow-up. It is important to note that women of color as only 17% (N = 7) of the sample at the eighteen-month follow-up were significantly underrepresented, which further limits the generalizability of the findings.

As in most randomized field experiments, this sample was not randomly selected, but included only those CSA survivors who volunteered
to participate in the Edmond et al.'s (1999) original study. Survivors who volunteer for such an experiment may be unlike those who do not volunteer. Consequently, participants in the present study may not have been representative of the population (i.e., adult female survivors of childhood sexual abuse). Therefore, the generalizability of the results may be limited. Despite these limitations, the present study was able to provide preliminary evidence for the therapeutic benefits of EMDR for adult female survivors of childhood sexual abuse, as well as for the ability of these benefits to be maintained long-term (i.e., over an 18-month follow-up period). Furthermore, the present study is unique in that it represents the only controlled study to date reporting the long-term effects of EMDR using a sample consisting exclusively of adult female CSA survivors. Consequently, additional controlled experiments are needed for more conclusive evidence as to the immediate and long-term benefits of EMDR for adult female survivors of childhood sexual abuse, as well as for other specified trauma populations. Ideally, such future investigations should have larger sample sizes, provide longer-term treatment with more than six sessions of EMDR before post-testing, and seek ethical ways to minimize participants utilizing subsequent treatment between post-testing and follow-ups.

The ability of the present study to truly assess the effects of Edmond et al.'s (1999) initial treatment conditions at 18 months was seriously impaired by multiple treatment interference, attrition and a lack of power. Despite these limitations, however, the data was able to provide some valuable information to help further our knowledge about the treatment needs of adult female survivors of childhood sexual abuse and the long-term effects of EMDR. The frequency with which the survivors obtained subsequent therapy after posttest, in conjunction with the fact that 90% of the original sample had participated in therapy prior to entering the study, supports the need for ongoing, possibly long-term, therapy for this population. This point is particularly relevant in this day of managed care, where the number of individual therapy sessions is often quite restricted. Brief symptom oriented treatment models are alluring, and while potentially helpful in alleviating symptom distress, they may be insufficient to address the myriad of issues with which survivors often struggle.

One must question, then, the value of offering an intervention that is as narrowly focused as the one initially evaluated by Edmond et al. (1999), which focused primarily on the one traumatic memory determined to be the most distressing by each participant. How useful can it be to survivors to focus solely on one issue? Despite the apparent limitations of such a
therapeutic approach, results of the present study did suggest that it was possible to achieve and maintain some positive gains from such a relatively short period of treatment. Reduction or alleviation of traumatic issues designated as most distressful by the participants and their accompanying symptoms, particularly those as disruptive as anxiety, depression, and PTSD, is no small matter. This is a particularly salient point for survivors with limited treatment options available to them due to a lack of insurance coverage or financial ability to pay for therapy.

It is also important to acknowledge that the elimination of such symptoms does not necessarily represent complete resolution of a childhood trauma. In reality, practitioners working with adult survivors, including those who use and those who do not use EMDR, are not likely to adopt such a narrowly focused treatment approach. Six sessions focused on one treatment issue was an artifact of an unfunded research project in Edmond et al.’s (1999) study. Despite the often-dramatic claims associated with EMDR (Shapiro, 1989; Lohr et al., 1995; Herbert & Mueser, 1992), practitioners who choose to use it with this population should bear in mind the likely need for longer-term work and should ensure that the CSA survivors seen in treatment are aware of this likelihood as well.

REFERENCES


