# The need for careful study design when investigating the benefits of psychological interventions for trauma survivors: a commentary on Renner, Bänninger-Huber, Peltzer (2011)

*Tracy M. McGuire*
*Christopher W. Lee*
*Peter D. Drummond*

School of Psychology, Murdoch University, Perth, Australia



**Author correspondence:**
Tracy McGuire
School of Psychology
Murdoch University
South Street, Murdoch
Perth, Western Australia 6150
Telephone: (61-08) 9360 2186.
Email: t.mcguire@murdoch.edu.au

URL: http://trauma.massey.ac.nz/issues/2014-1/AJDTS_18-1_McGuire.pdf

## Abstract

*Methodological concerns are explored and questions raised about the validity of conclusions reached in a recent article by Renner, Bänninger-Huber and Peltzer (2011). These authors reported treatment outcomes of Chechen asylum seekers and refugees with Posttraumatic Stress Disorder (PTSD), anxiety and depression following treatment with Group Cognitive Behavioural Therapy (CBT); a Culture-Sensitive and Resource Oriented Peer Group (CROP); Eye Movement Desensitization and Reprocessing (EMDR); and a wait-list condition. They concluded that CROP was significantly superior to wait-list and as effective as CBT in reducing symptomatology, and that EMDR was ineffective. However the study contains serious methodological problems including a lack of randomization information, a lack of independent evaluators, inadequate treatment fidelity, and inadequate treatment dosage. Furthermore, the small sample size, high attrition rate and unequal group numbers compromise the statistical power of this study, and possibly compromise the underlying statistical assumptions rendering any conclusions unreliable. This is serious given that misrepresentation of data is damaging to treatment models and clinical practice where such articles guide clinician's treatment choices.*

**Keywords:** *methodological rigour, research design, Posttraumatic Stress Disorder, EMDR*

Investigation into the effectiveness of treatments for Posttraumatic Stress Disorder (PTSD) in western and non-western cultures is vital as our societies become increasingly diverse. It cannot be assumed that a treatment proven to be effective in one culture will necessarily be effective when working with individuals from another culture.

An article published in this journal by Renner, Bänninger-Huber and Pelzer (2011) claimed to provide scientific data on the effectiveness of a Culture-Sensitive and Resource Oriented Peer (CROP) group method in comparison to Cognitive Behavioural Therapy (CBT), Eye Movement Desensitization and Reprocessing (EMDR) and a wait-list group. The participants were Chechen asylum seekers and refugees with symptoms of Posttraumatic Stress Disorder (PTSD). Methodological flaws such as unequal group sizes (CROP n=9, CBT n=10, EMDR n=6, WL n=7), unequal session numbers (CROP and CBT groups receiving 15 sessions in comparison to three EMDR sessions) and possible treatment and therapist bias compromised the results and the ability to compare EMDR to the other intervention groups. A critique of these methodological flaws is presented to highlight the importance of methodological rigour in treatment outcome studies.

Methodological differences in studies have been found to lead to different conclusions about treatment efficacy (Kazdin, 1994). In recent years it has been highlighted that the quality of reporting of randomised controlled trials (RCTs) is not optimal and that without transparent reporting, readers cannot judge the reliability and validity of trial findings.

A group of scientists and editors developed the CONSORT (Consolidated Standards of Reporting Trials) statement to improve the quality of reporting of RCTs (Schulz, Altman & Moher, 2010). Two notable studies specific to PTSD have also aimed to guide methodological rigour in this field; Foa and Meadows (1997) and Maxfield and Hyer (2002). Both referred to a gold standard research design which is a representation of standards that are understood when conducting and reviewing research in psychotherapy. These guidelines have enhanced our capacity to design, evaluate and draw accurate conclusions that ultimately guide our clinical decisions.

Foa and Meadows (1997) described seven parameters as being critical to a methodologically strong outcome study: 1) Clearly Defined Target Symptoms; 2) Reliable and Valid Measures; 3) Use of Blind Evaluators; 4) Assessor Training; 5) Manualized, Replicable, Specific Treatment Programs; 6) Unbiased Assignment to Treatment; 7) Treatment Adherence. Maxfield and Hyer (2002) looked at whether differences in research outcomes were related to methodological differences. Their study employed a gold standard research scale, adapted from Foa and Meadows (1997), to critique methodological strengths and weaknesses and their association with effect sizes for research publications on EMDR. This examination demonstrated a significant correlation between gold standard research methods and treatment outcomes. Maxfield and Hyer (2002) concluded that assessment reliability and treatment fidelity were critical factors in methodological rigour. Overall, as the methodology became more rigorous, the treatment effect had become positive and size of the effect larger (Maxfield & Hyer, 2002). Bearing these results in mind, we have used these gold standards to critique the paper by Renner, Bänninger-Huber and Pelzer (2011).

The lack of adequate information regarding randomization in the Renner et al. (2011) paper is of major concern. Although the authors stated that "participants were assigned to the above mentioned conditions at random" (p.5), the process of random assignment was not disclosed. The unequal cell sizes (CROP n=25, CBT n=21, EMDR n=17 and Wait List (WL) n=31) may indicate a lack of adequate random allocation sequencing. According to CONSORT guidelines, randomization procedures and the allocation ratio should be identified as it is an integral part of controlled research (Schulz et al., 2010). Allowing readers to assess the generation of the random allocation sequence permits them to uncover the likelihood of selection bias in group assignment; and to assess whether any differences in outcomes between groups reflect the treatment rather than extraneous factors (Foa & Meadows, 1997; Maxfield & Hyer, 2002; Schulz, Chalmers, Hayes & Altman, 1995).

To ensure that therapist and treatment effects can be separated, not only should the allocation to treatment condition be randomized but also treatment should be delivered by at least two therapists to whom participants are randomly assigned (Foa & Meadows, 1997; Maxfield & Hyer, 2002). In the Renner, Bänninger-Huber and Pelzer (2011) study participants in the CBT and CROP groups were assigned to one of two therapists based on gender, whereas just one therapist delivered EMDR. This design introduced the possibility of therapist effects, where therapist characteristics such as training and competence, personal characteristics and experience levels can interfere with treatment delivery and treatment outcomes (Elkin, 1999). The introduction of a second therapist in the EMDR group, and computer generated randomization to conditions and therapists would have removed extraneous factors while also ensuring equal distribution to each condition.

Gold standard research requires use of blind independent assessors to combat expectancy and demand bias in participants and therapists. Renner, Bänninger-Huber and Pelzer (2011) did not disclose whether the assigner of conditions was blind to participant assessment or whether the evaluator (the first author) was blind to condition allocation when collecting outcome data. This introduced the possibility of bias in the study results. Furthermore, there was no disclosure stating whether those assessing the outcome data were blind. If they were not there could have been a bias in the selection of analytical strategies and removal of data or selection of time points (Wood, Egger, Gluud, Schulz, Juni, & Altman et al., 2008). Ultimately, this lack of clarity raises questions about the validity of the study results.

The main goal in a treatment outcome study is the specification of treatments and an evaluation of their feasibility and efficacy (Perepletchikova, Treat & Kazdin, 2007). The interpretation of treatment effects requires affirmation that the treatment was delivered as it is designed. Otherwise ambiguity in evaluating both what the intervention was and why it produced effects is introduced (Kazdin, 2003; Perepletchikova, Treat & Kazdin, 2007). Lack of treatment protocol and treatment fidelity checks introduces possible inconsistencies and bias in treatment delivery across patients and therapists. Maxfield and Hyer (2002) found a "large significant correlation between treatment fidelity and effect size" (p.36) when reviewing EMDR studies. They reported that studies which assessed treatment fidelity showed larger treatment outcomes than studies that did not assess treatment fidelity. The Renner et al. (2011) paper used three treatment programs, CBT, CROP and EMDR, in their study and did not appear to follow specific treatment protocols, nor were there any reported treatment fidelity checks to ensure treatment integrity. These deficiencies not only introduced possible inconsistencies and bias but also compromised the replicability of the study. Incorporating a treatment protocol for EMDR and CBT, a documented manual for CROP groups and a sufficiently experienced independent evaluator of treatment delivery would have eliminated confounds and possible bias.

Methodological questions must also be raised in relation to the sample size and attrition rates. The section of the Renner et al. (2011) paper outlining the design specifies data were collected from 94 participants (CROP n=25, CBT n=21, EMDR n=17 and Wait List (WL) n=31). Within the paper it was stated that analysis of data was completed on only 32 participants (CROP n=9, CBT n=10, EMDR n=6 and WL n=7). A low N can decrease the statistical power therefore influencing the likelihood of a type I error (i.e., concluding the means were different when they were not hence recording a false positive result); or type II error (i.e., concluding the means were not different when there is a difference hence recording a false negative result). Such errors can be addressed by assessing the adequacy of the statistical power when choosing the statistical analysis and the alpha level to be used. The authors stated that the "sample size was planned beforehand with respect to expected statistical power" (p.7), indicating the consideration of statistical analysis and therefore statistical power for a sample size of 94. There was no indication of an adjustment to the type of analysis or the alpha level, using techniques such as the Bonferroni correction, in order to conserve statistical power (Gravetter & Wallnau, 2007).

The high attrition rates raise the possibility that key assumptions required for statistical analysis may be violated (Gravetter & Wallnau, 2007). CONSORT guidelines specify that losses and exclusions following random allocation should be reported and discussed. Without this it is difficult to determine the reason for attrition and whether the attrition affects the interpretation of results (Schulz et al. 2010). In the paper the authors did not specify at what phase of the study and why participants were lost. An intention to treat analysis could have been used to avoid erroneous conclusions (Wood, White & Thompson, 2004). Ultimately the study seems underpowered statistically and the authors did not identify changes to methods or statistical considerations to combat this decrease in statistical power. Interpretation of the results section of Renner et al. (2011) is also difficult. The authors reported N and p values but did not include F ratios, degrees of freedom, effect size or confidence intervals. These flaws make it difficult to draw significant conclusions from the study data and in identifying any clinically meaningful treatment effects.

Of great concern is the inadequacy of treatment dosage, treatment fidelity and sample size for the EMDR condition. The final sample size in the EMDR group was only six participants. Furthermore, Renner et al. (2011) specified that only 50% of the EMDR group received the "actual EMDR technique" (p.1). EMDR treatment was defined in the paper as consisting of three sessions during which the therapist collected participants' trauma history; conducted brief counselling; and, if the therapist felt that it was appropriate, administered the actual EMDR technique. Thus, three participants appear to have received only 1 to 2 sessions of an EMDR technique. No treatment protocol or independent fidelity check was used to clarify what the EMDR technique consisted of, introducing possible bias and treatment confounds that compromise outcomes.

Although further methodologically rigorous studies are required to ascertain the ideal treatment dosage of EMDR in multiple trauma populations, there is empirical literature (mainly pertaining to veterans and child abuse victims) that suggests this population requires additional sessions (Carlson, Chemtob, Rusnak, Hedlund & Muraoka, 1998; Boudewyns & Hyer, 1996). For example, Russell, Silver, Rogers and Darnell (2007) concluded that combat veterans who had suffered multiple traumas required between 3.8 (nonwounded) and 8.5 (wounded) sessions of EMDR. Maxfield and Hyer (2002) suggested that an adequate course of treatment be included as part of the gold standards of research, as their findings indicate an insufficient course of EMDR treatment may interfere with treatment efficacy. It has been noted by the International Society for Traumatic Stress Studies (Chemtob, Tolin, vander Kolk & Pitman, 2000) and the Departments of Veterans Affairs and Defence Joint Clinical Practice Guidelines for PTSD (see Russell et al., 2004) that randomized studies of EMDR, with veterans who have experienced multiple traumas, are often flawed due to insufficient treatment doses for this population. Thus, one or two sessions of EMDR most likely is insufficient.

This leads to questions relating to the authors' level of understanding regarding the theory and practice of EMDR. Renner, et al. (2011) reported that the EMDR technique was not used in some cases due to the inability of the participant to visualize a specific traumatic event, because they had incurred multiple traumas.

However, existing research has been conducted consistently demonstrating that EMDR is beneficial for clients with multiple traumas. Spates, Koch, Cusack, Pagoto and Waller (2008) indicated in their meta-analysis that EMDR is efficacious in treating both civilian and military populations who often incur multiple traumas. Carlson et al. (1998) randomly assigned 35 Vietnam combat veterans with PTSD to 12 sessions of EMDR, biofeedback relaxation (RXT) or a control group. The Results of this study showed only 2 of 9 participants in the EMDR group met criteria for PTSD on the CAPS at follow-up, a significantly greater reduction than the RXT group where 7 of the 9 participants met criteria for PTSD at follow-up. Boudewyns and Hyer (1996) compared 61 veterans who received 5 to 7 sessions of EMDR with eyes open and eyes closed plus 8 group sessions with treatment as usual plus group sessions. All groups improved significantly on structured interviews measuring PTSD symptoms; the two EMDR groups (with and without eye movements) showed larger improvements on mood and physiological measures compared to standard therapy. Edmund, Rubin and Wambach (1999) used a randomized control study to assess the efficacy of EMDR with 59 adult female childhood sexual abuse survivors. Following six 90 minute EMDR sessions symptoms decreased significantly more than in the control group. Rothbaum (1997) randomly assigned 18 adult female rape victims, most suffering from multiple traumas, to four 90 minute sessions of EMDR compared to a wait list control. Scores on PTSD and depression scales decreased significantly in the EMDR group, with 90% no longer meeting full criteria for PTSD after treatment.

In sum, the Renner et al.(2011) study contained methodological flaws that reduce confidence in the overall conclusions about the effectiveness of EMDR, CBT and CROP treatments. All research has flaws. The question is to what extent the flaws erode the credibility of the research. In this case we suggest the answer is ‘substantially’, due to failure to meet several gold standard criteria: randomization; blind evaluators; treatment adherence / fidelity; and treatment dosage. This applied particularly to EMDR but also raises questions about the strength of treatment effects for CBT and CROP in this study of trauma survivors.

## References

Boudewyns, P.A., & Hyer, L.A. (1996). Eye movement desensitization and reprocessing (EMDR) as treatment for post-traumatic stress disorder (PTSD). *Clinical Psychology and Psychotherapy, 3*, 185-195.

Carlson, J.G., Chemtob, C.M., Rusnak, K., Hedlund, N.L. & Muraoka, M.Y. (1998). Eye movement desensitization and reprocessing for combat-related posttraumatic stress disorder. *Journal of Traumatic Stress, 11*, 3-24.

Chemtob, C., Tolin, D., van der Kolk, B.A. & Pitman, R. (2000). Eye movement desensitization and reprocessing. In E.B. Foa, T.M.Keane, & M. J. Friedman (Ed.s), *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies* (pp. 333-335). New York, NY: Guilford Press.

Elkin, I. (1999). A Major Dilemma in Psychotherapy Outcome Research: Disentangling Therapists From Therapies. *Clinical Psychology : Science and Practice, 6,* 10-32.

Foa, E.B. & Meadows, E.A. (1997). Psychosocial treatments for posttraumatic stress disorder: A critical review. *Annual Review of Psychology, 48*, 449-480.

Gravetter, F.J. & Wallnau, L.B. (2007*). Statistics for the behavioral sciences* (7th ed.). Belmont, CA: Thompson Wadsworth.

Kazdin, A.E. (1994). *Handbook of psychotherapy and behaviour change* (4th ed.). Oxford, UK: John Wiley & Sons.

Kazdin, A.E. (2003). *Research design in clinical psychology* (4th ed.). Boston, IL: Allyn & Bacon.

Kazdin, A.E. & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. *Journal of Consulting and Clinical Psychology, 57*, 138-147.

Maxfield, L. & Hyer, L. (2002). The Relationship between efficacy and methodology in studies investigating EMDR treatment of PTSD. *Journal of Clinical Psychology, 58*, 23-41.

Perepletchikova, F., Treat, T.A., & Kazdin, A.E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*, 829-841.

Renner, W., Bänninger-Huber, E. & Peltzer, K. (2011). Culture-sensitive and resource oriented peer (CROP) -groups as a community based intervention for trauma survivors: A randomized controlled pilot study with refugees and asylum seekers from Chechnya. *The Australasian Journal of Disaster and Trauma Studies, 2011,* 1.

Rothbaum, B.O. (1997). A controlled study of eye movement desensitization and reprocessing in the treatment of post-traumatic stress disordered sexual assault victims. *Bulletin of the Menninger Clinic, 61*, 317-334.

Russell, M.C., Silver, S. M., Rogers. S., & Darnell, J. (2007). Responding to an identified need: A joint Department of Defense-Department of Veterans Affairs training program in eye movement desensitization and reprocessing (EMDR) for clinicians providing trauma services. *International Journal of Stress Management, 14*, 61-71.

Schulz, K.F., Altman, D.G., & Moher, D., for the CONSORT Group (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Annals of Internal Medicine, 152*(11), 1-8.

Schulz, K.F., Chalmers, I., Hayes, R.J. & Altman, D.G. (1995). Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association, 273*, 408-412.

Spates, C.R., Koch, E., Cusack, K., Pagoto, S. & Waller, S. (2008). Eye movement desensitization and reprocessing. In E.B. Foa, T.M. Keane, M. Terence, M.J. Friedman, & J.A. Cohen, (Ed.s), *Traumatic stress* (pp.279-305). New York, NY: Guilford Press.

Wood, L., Egger, M., Gluud, L.L., Schulz, K.F., Juni, P., Altman, D.G., Gluud, C., Martin, R.M., Wood, A.J.G., & Sterne, J.A.C. (2008). Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: Meta-epidemiological study. *British Medical Journal, 336*, 601-605.

Wood, A.M., White, I.R., & Thompson, S.G. (2004). Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clinical Trials, 1*, 365-376.

*This page intentionally left blank.*