What does the data say about the importance of eye movement in EMDR?

Christopher William Lee, Pim Cuijpers

PII: S0005-7916(13)00076-1
DOI: 10.1016/j.jbtep.2013.10.002
Reference: BTEP 984

To appear in: Journal of Behavior Therapy and Experimental Psychiatry

Received Date: 1 October 2013
Accepted Date: 3 October 2013


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
What does the data say about the importance of eye movement in EMDR?

Christopher William Lee, Pim Cuijpers
a School of Psychology and Exercise Science, Murdoch University, South St., Murdoch WA 6150, Australia
b Department of Clinical Psychology and EMGO Institute for Health and Care Research, VU University and VU University Medical Center Amsterdam, The Netherlands

Earlier this year we published a paper that gave an up-to-date review of the evidence for whether eye movement had an effect in facilitating the processing of trauma memories (Lee & Cuijpers, 2013). In this paper we did a meta-analysis to look at this issue both in laboratory contexts and in treatment studies that used Eye Movement Desensitisation and Reprocessing (EMDR). Devilly, Lohr and Ono (in press) have provided a commentary on our article that contains many inaccuracies and several irrelevant points that could have the effect of clouding our findings. We therefore take this opportunity to repeat the main findings which were not addressed in their commentary and document both the irrelevancies and inaccuracies.

The first main finding of the meta-analysis, which is not discussed anywhere in the commentary, was a large and significant (d=.74) difference in effect size in favour of eye movements over no eye movements in the laboratory studies. Another finding for these participants was that pairing eye movement with their trauma memories led to particularly large and significant reductions on measures of vividness (d=.91) compared to no eye movements while focusing on trauma memories. This is an important practical point. Clients often don’t want to relive their experiences as required for traditional exposure treatments and those with high symptom levels respond poorly to such treatments (Scott & Stradling, 1997). Since EMDR does not require this reliving process (Lee, Taylor, & Drummond, 2006), the distancing effect produced by eye movements provides such clients with an opportunity to focus on the trauma memories without being overwhelmed.
Secondly, the effect size for the treatment studies was misrepresented. In the commentary the .27 significant effect size was quoted from our findings and a discussion on how the associated confidence interval was close to zero followed. In fact the main finding from the treatment studies was an effect size of .41 from the 15 study trials. This is not small but within a moderate effect size range and it is significant. Two studies in this sample appeared to contribute most to the heterogeneity. Both had strong effect sizes in favour of eye movements. When these studies were removed, the sample was extremely homogenous ($I^2 = 0$) and the difference in effect size for eye movement remained significant (although reduced to .27). Therefore, no matter how one looks at this data there is a significant effect for eye movement.

The commentary contains a number of irrelevant and/or inaccurate comments which are a distraction from these main findings. These will be discussed below.

The first inaccurate comment they make concerns alleged missing studies from the analysis. The studies that they cite as examples were mostly simply irrelevant to the scientific question we sought to answer. That is under identical randomised conditions, is there value in adding an eye movement component. We did not seek to do an analysis to assess the relative efficacy of EMDR versus traditional exposure. Hence studies that directly compared these disparate treatments such as (Devilly & Spence, 1999; Taylor et al., 2003) were not included in our analysis. Numerous other studies also did this (de Roos et al., 2011; Ironson, Freud, Strauss, & Williams, 2002; Lee, Gavriel, Drummond, Richards, & Greenwald, 2002; Nijdam, Gersons, Reitsma, de Jongh, & Olff, 2012; Rothbaum, Astin, & Marsteller, 2005). However the manuals for the treatments used in these studies were not identical. There are dramatic
differences in doing traditional exposure compared to EMDR such as the therapist directing the client to relive the experience and not permitting the client to deviate from the targeted trauma (Lee, 2008). In contrast the studies we included all involved the therapist giving the same instruction in both conditions so that the pure effect of eye movement can be assessed. Other studies we excluded did not have a randomisation process or order effects were not controlled for (Montgomery & Ayllon, 1994). In that study eye movements were reported to have a positive effect but given this failure in randomisation we excluded it, even though it was included in a previous meta-analysis (Davidson & Parker, 2001).

Our paper was also criticized for not citing a prior meta-analysis by van Etten and Taylor (1998). However this was also simply not relevant. van Etten and Taylor never calculated an effect size for eye movement compared to no eye movement. Instead, they examined published PTSD treatment studies that contained medication, psychotherapy, EMDR, and traditional exposure. That particular meta-analysis had no comment to make about data on the role of eye movements other than to say there were not sufficient trials to comment on this issue. To cite this as an omission is puzzling as the article was positive about the use of EMDR in that they noted that EMDR achieved similar effect size to other treatments but with fewer sessions.

Another major irrelevant, and in this case inaccurate, criticism in the commentary was alleging that we counted studies twice. It is common in meta-analysis, where a treatment is investigated under different conditions in the same study, that these are counted as separate trials. This is not counting the same participant twice if each participant is not part of each trial. It enables each participant within the trial to be compared to other participants under identical conditions. An example of this is when eye movements are compared to no eye
movements with participants instructed to relive their experiences and there is also data on another set of subjects who are treated with eye movements or no eye movements while being asked to stay distant from their trauma (Lee & Drummond, 2008). A meta-analysis cited in the commentary (van Etten & Taylor, 1998) also followed this tradition and analysed 41 treatment studies which they described contained 68 trials. Even if one was to pool disparate studies to obtain a single effect size calculation, which is implied in the commentary, theoretically the effect size would be similar but the degrees of freedom would go down rather than up (as there is one less study). Thus the p-value for our obtained effect sizes would have been larger than that reported. As final proof of the irrelevancy of the Devilly et al. argument is that we conducted two sensitivity analyses in which we included only one of the two comparisons from Lee & Drummond, one in which the highest effect size was removed from the analyses, and one in which the lowest effect size was removed. As can be seen in Table 2 of our paper, these sensitivity analyses showed that removal of any one comparison hardly had any effect on the overall results.

The commentary also alleges that we were not clear on how our calculations of significance and effect sizes were made. They state "We assume that the interval around this average was computed based upon total sample size rather than the number of measures combined or a multiplication of number of measures by sample size or some other estimate of error". We have no idea how they made this assumption or how there is any doubt on our methods. The methods on how we calculated effect sizes are clearly described on page 232 and 233. For example on page 233 we described how multiple measures were handled. When multiple measures were used, the resulting data were first pooled within the study, and the pooled result was used in the meta-analysis. This is currently the best method to handle this issue,
and this is what the state-of-the-art software we used (Comprehensive Meta-analysis: CMA) does.

One comment that has potential merit was that we included in the treatment studies, trials where participants that did not have a DSM diagnosis with trials where there was a DSM diagnosis. The implication is that perhaps eye movements delivered as part of a treatment package may function differently in these populations. We decided to combine these studies and test empirically whether that was the case. In the subgroup analysis, we found that there was no significant difference in effect size difference for eye movements whether or not there was a DSM diagnosis. Similarly, we found that there was no significant difference in effect size difference for eye movements whether the sample treated trauma memories of students or people presenting to a clinic (Table 2 in Lee & Cuijpers, 2013). Although a larger sample of studies is needed to be confident in generalising from these findings.

Another irrelevant criticism of our paper was that we included SUDS values that were available at end of treatment as a measure of treatment effectiveness. We reject that this is inappropriate. SUDS can be both an outcome and process measure. In traditional exposure, SUDS are used during the session to assess how the habituation process is proceeding and to help ascertain ‘hot spots’ which are the subject of further attention by the therapist. As said, SUDS can also be an outcome measure. At the conclusion of treatment, if the process is successful, then there should be no hot spots and the therapist will check to see that SUDS are low. SUDS are also used in EMDR to check the current degree of distress to the memory. Successful outcome of any PTSD treatment is a reduction in the frequency of avoidance and intrusive symptoms and that when a person is reminded of the trauma that it is not accompanied by hyperarousal. SUDS can help to assess this last aspect of recovery.
Certainly clients haunted by trauma memories are pleased when they can think of the event and no longer experience distress. Devilly and colleagues’ criticism of this issue is also moot given that when we excluded SUDS and VOC scores from the treatment studies, the effect size for eye movement remained significant (see Table 2).

A further inaccuracy in the commentary of Devilly et al. (in press), was in the citing of Z scores for whether training or use of a manual moderated the effect size of eye movement. For example for the analysis of treatment manual, they cite a Z score of .73 with a p-value of .46. This is simply incorrect. The actual data presented in Table 2 indicated the effect size for eye movements over no eye movements was 3.21 for treatment studies that used a published manual and near zero for those that did not. The p-value of the effect size difference between the studies was .03.

Another inaccurate comment was that it is asserted that the Shapiro study should not have been included in the sub-analysis of whether a manual was used. It was not included.

The second last paragraph of the commentary contains a hypothesis that the treatment effects due to eye movements can be explained by distraction and that as such the effect of EMDR weakens over time. This is simply not supported by any meta-analysis data (Bisson & Andrew, 2007; Ho & Lee, 2012; Rodenburg, Benjamin, de Roos, Meijer, & Stams, 2009).

We stand by our criticism of a previous meta-analysis (Devilly, 2002) as being open to bias. Guidelines on reporting meta-analysis findings have for the last 13 years emphasised that more than one rater should be used to classify and calculate the effect size (Stroup et al., 2000). This reduces the risk of possible rater-bias. Hence the Meta-analysis Observational
Studies in Epidemiology (MOOSE) Group published a checklist on good practice which includes this point (Stroup et al., 2000). A strength of our study was that multiple raters were used.

In conclusion, the commentary by Devilly and colleagues not only ignored the most significant findings in our meta-analysis but is clouded by red herrings and untruths. These included descriptions of the first author as a reseller, criticism of the use of CMA software, listing irrelevant studies, an inaccurate portrayal that we counted certain studies twice, and clear factual errors. These inaccuracies were corrected above so that the data from Lee and Cuijpers (2013) can be interpreted from a more clear perspective.

References


