Journal of Behavior Therapy and Experimental Psychiatry 63 (2019) 48–56

Contents lists available at ScienceDirect

Journal of Behavior Therapy and Experimental Psychiatry

journal homepage: www.elsevier.com/locate/jbtep

Effects of "Visual Schema Displacement Therapy" (VSDT), an abbreviated EMDR protocol and a control condition on emotionality and vividness of aversive memories: Two critical analogue studies

Suzy J.M.A. Matthijssen[a,b,*], Lucinda M. van Beerschoten[c,d], Ad de Jongh[e,f,g,h], Irene G. Klugkist[i,j], Marcel A. van den Hout[a,b]

[a] Altrecht Academic Anxiety Centre, Altrecht GGz, Utrecht, the Netherlands
[b] Department of Clinical Psychology, Utrecht University, Utrecht, the Netherlands
[c] Stichting Elisabeth, Breda, the Netherlands
[d] Careyn, Maarssen, the Netherlands
[e] PSYTREC, Bilthoven, the Netherlands
[f] Academic Centre for Dentistry Amsterdam (ACTA), University of Amsterdam and VU University Amsterdam, the Netherlands
[g] School of Health Sciences, Salford University, Manchester, United Kingdom
[h] Institute of Health and Society, University of Worcester, United Kingdom
[i] Methodology and Statistics of Social and Behavioural Sciences, Utrecht University, Utrecht, the Netherlands
[j] Research Methodology, Measurement and Data Analysis of Behavioural, Management and Social Sciences, Twente University, Enschede, the Netherlands

ARTICLE INFO

Keywords:
VSDT
Visual Schema Displacement therapy
EMDR
Trauma therapy
PTSD

ABSTRACT

Background and objectives: Visual Schema Displacement Therapy (VSDT) is a novel therapy *which has been described as a treatment for* stress and dysfunction caused by a traumatic event. Although its developers claim this therapy is quicker and more beneficial than other forms of trauma therapy, its effectiveness has not been tested.
Methods: We compared the efficacy of VSDT to an abbreviated EMDR protocol and a non-active control condition (CC) in two studies. In Study 1 participants ($N = 30$) were asked to recall three negative emotional memories under three conditions: VSDT, EMDR, and a CC, each lasting 8 min. Emotional disturbance and vividness of the memories were rated before and after the (within group) conditions. The experiment was replicated using a between group study. In Study 2 participants ($N = 75$) were assigned to one of the three conditions, and a follow-up after 6–8 days was added.
Results: In both studies VSDT and EMDR were superior to the CC in reducing emotional disturbance, and VSDT was superior to EMDR. VSDT and EMDR outperformed the CC in terms of reducing vividness.
Limitation: Results need to be replicated in clinical samples.
Conclusions: It is unclear how VSDT yields positive effects, but irrespective of its causal mechanisms, VSDT warrants clinical exploration.

1. Introduction

EMDR therapy, a psychological treatment for PTSD, was introduced in 1989 (Shapiro, 1989) and was received with skepticism at first (e.g. Herbert et al., 2000; Lohr, Kleinknecht, Tolin, & Barrett, 1995). For example, according to Lileinfeld (1996) the continued widespread use of EMDR therapy for therapeutic purposes was “only another example of human mind's willingness to sacrifice critical thinking for wishful thinking.” However, over the past decades several meta-analyses have shown EMDR to be highly effective in the treatment of PTSD (Bisson et al., 2007; Bradley, Greene, Russ, Dutra, & Westen, 2005; Chen et al., 2014; Lee & Cuijpers, 2013; Seidler & Wagner, 2006) and, as a result, EMDR has been recommended as a first-line treatment for PTSD in treatment guidelines in the Netherlands, United Kingdom, France, United States, Israel, Northern Ireland and by the World Health Organization (American Psychiatric Association, 2004; Bleich, Kotler, Kutz, & Shalev, 2002; CREST, 2003; Dutch National Steering Committee Guidelines Mental Health Care, 2003; Haute Autorité de la Santé, 2007; National Institute for Clinical Excellence, 2005; World Health Organization, 2013).

* Corresponding author. Altrecht GGZ, Nieuwe Houtenseweg 12, 3524 SH, Utrecht, the Netherlands.
E-mail address: s.matthijssen@altrecht.nl (S.J.M.A. Matthijssen).

https://doi.org/10.1016/j.jbtep.2018.11.006
Received 18 April 2018; Received in revised form 9 November 2018; Accepted 25 November 2018
Available online 27 November 2018
0005-7916/

A crucial element in EMDR therapy is the recall of emotionally disturbing memories while engaging in eye movements (EM), typically induced by following the therapist's fingers moving the hand back and forth in front of the patient's eyes. Lee and Cuijpers (2013) conducted two separate meta-analyses, one including laboratory and the other clinical studies, showing the effectiveness of employing EM while holding a memory in mind with medium to large effect sizes in laboratory studies, and medium effect sizes in clinical studies.

The claim that Recall + EM is effective in reducing the emotionality of disturbing memories was not based on earlier scientific studies and was unconventional. An even more unconventional claim, not derived from science either, has been made by Nik and Eva Speakman. They claim that emotionally disturbing memories can be successfully processed by instructing patients to follow a watch with their eyes. More specifically, the therapist draws a circle with a watch in front of the patient, and asks the patient to bring up an image of a person or a memory of an event which made them laugh, and indicate where in the circle - while recalling this mental representation and following the watch –they feel the urge to laugh the most. This procedure is repeated for the emotionally disturbing memory. The therapist then moves the watch quickly from the most disturbing point (the 'trauma point') to the point where the patient feels the urge to laugh (the 'laughter point') while saying 'Whoosh!' out loud. The patient is instructed to keep their eyes focused on the watch, blink and sigh (see 'Procedure' for a more elaborative description of the Visual Schema Displacement Therapy, or VSDT, formally called Visual Coding.

This claim about its effectiveness has many elements of pseudoscience, particularly the claim of being a miracle cure for a wide array of mental disorders (e.g., PTSD, obsessive compulsive disorder, agoraphobia, specific phobia) and with an unclear theory about the working mechanism (Cuijpers & Cristea, 2016). More importantly, there is complete lack of empirical evidence concerning the efficacy of this therapy. Yet, as Meyers (2007) pointed out "a large number of significant discoveries in medicine arose, and entirely new domains of knowledge and practice were opened up, not as a result of painstaking experimentation, but rather from chance and even outright error." Some accidental discoveries (e.g. X-rays, the first antibiotic, chemotherapy drugs and Dramamine as a drug in preventing and relieving motion sickness) stem from observations that were curious and may very well have been unreliable (Meyers, 2007). This warrants keeping an open-mind about the extraordinary claims, such as made in the case of VSDT.

The authors, decided to put the procedure to a critical test to assess its effectiveness in reducing the disturbance of emotional memories. As mentioned, the Recall + EM versus a Control Condition (CC) procedure has been a well-accepted paradigm to study (e.g. Andrade, Kavanagh, & Baddeley, 1997; De Jongh, Ernst, Marques, & Hornsveld, 2013; Maxfield, Melnyk, & Hayman, 2008; Wilson, Becker, & Tinker, 1995). An even more stringent condition was used here. That is, we decided to test not only whether VSDT was superior to a non-active control condition (CC), but also whether VSDT effects could be similar or even superior to the effects observed in laboratory models of EMDR,. Two experiments were conducted: The first used a within-group design comparing recall + EM with VSDT, and a CC; the second experiment employed a more stringent between-group design to allow for assessment of follow-up results.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Data from 30 participants (22 female, 8 male) from Utrecht University ($M = 22.53$, $SD = 3.69$) were included in the analyses. Forty-four participants, who received either course credits or financial compensation for participating, were recruited initially, but six participants did not complete the procedure; five were unable to select three negative memories, and one appeared not to be a student, which was an inclusion criterion. Eight further participants were excluded after completing the procedure (see Results).

2.1.2. Materials

Subjective intensity of disturbance: Subjective intensity of disturbance or distress felt by the individual due to an image or an emotional memory being recalled was indexed by a 11-point Likert-type subjective anxiety scale, the Subjective Units of Disturbance (SUD) scale. The scale ranges from 0 ("none at all") to 10 ("maximum distress"). The SUD scale was introduced by Wolpe (1969) and incorporated in the standard EMDR protocol (Shapiro, 1995). In the present study, participants were requested to indicate the SUD score verbally to the research assistant at the start of and after each condition.

Vividness: Perceived intensity of vividness of an image or emotional memory being recalled was measured using vividness scores ranging from 0 ("not vivid at all") to 10 ("very vivid"). Participants indicated this score verbally to the research assistant at the start of each condition, and after each condition.

Subjective urge to laugh: The subjective urge to laugh felt by the individual attributed to a happy emotional memory or a person being recalled was indexed using a laughter-point-score. The laughter point had to be at ≥ 7 on a scale of 0 ("no urge to laugh") to 10 ("maximum urge to laugh") ($M = 8.33$, $SD = 0.88$). Participants indicated this score verbally to the research assistant at the beginning of the VSDT condition.

2.1.3. Procedure

Study procedures were approved by the Facultairy Medical Ethics Committee of the Utrecht University (The Netherlands) (FETC16-101). Students were recruited at Utrecht University. They were asked whether they were interested in participating in an experiment concerning emotional memories for either credits or money. The duration of the experiment was 50 min. The experiment was carried out by a research assistant, who was part of the author team and was trained in VSDT by the originators of VSDT. Training in the EMDR procedure was provided by an EMDR Europe accredited trainer. Fidelity checks were based upon video recordings that were carried out on a pilot sample to ensure the procedure was carried out properly. When participants agreed to participate, an appointment was made for a later date. Once the experiment started, they were given an information letter which informed them about the procedure of the experiment, anonymity procedures, the right to stop the experiment at any time without consequences, and included instructions to inform the research assistant if they were feeling too uncomfortable to recall the emotional memories. If the participants had no questions, an informed consent was given for signing. After signing, a screening checklist was completed which included questions about age, education level, sex, and exclusion criteria. Participants were excluded if they were not students, or if they had disrupting visual problems, which consisted of any visual problem which made it unable to see the research assistant's actions necessary for the experiment (e.g., uncorrected impairment, partial loss of vision). Participants were also excluded when they had a possible interaction between their attention or concentration and when they used medication, alcohol, drugs, had current psychiatric treatment, or suffered from fatigue or extreme stress. After the checklist, three emotional memories were selected and were rated as "somewhat disturbing", "reasonably disturbing", "fairly disturbing", "quite disturbing" or "very disturbing". Only memories rated as "fairly disturbing", "quite disturbing" or "very disturbing" were included. In total, 57.8% of the memories were rated as "fairly disturbing", 32.2% as "quite disturbing" and 10% as "very disturbing". The participants formulated a (few) keyword(s) per memory and they estimated the duration passed since the memory event had taken place. The research assistant asked every participant for a summary of the event to check for any irregularities (e.g. a

suspected fake memory or a psychotic episode). For the next procedure, they were asked to choose the keyword that best suited the memory.

Memories were ranked by emotional disturbance (by asking the participants to indicate the rank order). Counterbalancing was then applied for the memories (most, middle, least disturbing) and the conditions (VSDT, EMDR, CC). This resulted in 36 different sequences, to which participants were sequentially allocated. All three conditions took 8 min. In this time, the most disturbing part of the selected memory ("target") was recalled, rated on a scale of 0–10 on emotional disturbance and vividness, and then either EMDR, VSDT or the CC was applied, depending on the sequence.

The EMDR procedure consisted of an abbreviated version of the Dutch version of the EMDR Standard protocol (De Jongh & ten Broeke, 2012). In the EMDR Standard protocol patients are requested to describe a memory in global storylines and select the most disturbing image of the memory ("target") in present time. In the present study participants recalled a memory, but were not specifically asked for the global storylines. They were asked to select the most disturbing part of the memory (target) in the present time, which also included an image in this procedure. Similar to the EMDR protocol they were asked to rate the Subjective Units of Disturbance (SUD; see Materials) of the target and to point out the location in the body where this disturbance was felt the most. For this study, the participants were also asked to rate the vividness of the target. Then, while recalling the target, EM were evoked by the research assistant. She moved her fingers horizontally in front of the participant for sets of 30 s each, about 30 cm from the participants' eyes. After each set of EM, the research assistant asked for associations, and directly after she instructed the participants to concentrate on that what came to mind. When no associations occurred, or the participant mentioned the same association three times the SUD of the target was re-rated, after which the sets of 30 s EM were continued. After 8 min, the alarm went off and the participant was asked to re-rate the target on SUD and vividness.

In the VSDT condition participants were told which memory was being targeted, but before this, they were asked to select a mental representation of a person or a memory of an event which made them laugh. They had to select a keyword for this memory or person. The research assistant stood in front of the participant at approximately 1.25 m, while the participant was seated. The assistant held a watch with the clock face towards the participant and drew a circle with a diameter of about 1.5 m in a clockwise motion from the assistants' point of view. The participant was instructed to follow the watch with their eyes, and to indicate where in the circle the strongest urge to laugh was felt. This point was indicated as the "laughter point", which was given the name of the keyword. The participant calibrated the urge to laugh at the laughter point on a scale of 0 ("no urge to laugh") to 10 ("maximum urge to laugh"). Then the same procedure was repeated for the emotional memory. The participant was asked to select the most disturbing part of the memory ("target"), and to indicate where in the circle the "trauma point" – the point where he felt the most disturbance - was located. The participant rated the SUD and the vividness of the target. After that, the assistant explained the procedure. Then, she instructed the participant to keep his or her eyes focused on the watch and subsequently moved the watch quickly from the trauma point to the laughter point while saying out loud "Whoosh!". Next, the participant had to blink repeatedly, to squeeze the eyes tight and to make two deep sighs. After repeating the procedure three times participants rated by how much the SUD score associated with the target was reduced compared to the previous rating. This procedure was repeated. After 8 min, the alarm went off and the target was re-rated on SUD and vividness.

In the non-active control condition (CC), participants were asked to select the most disturbing part ("target") of the emotional memory, and rate its disturbance and vividness. Next, the participants were instructed to do nothing and relax and were told it did not matter what they thought about. After 8 min – when the alarm went off –the target was re-rated on SUD and vividness.

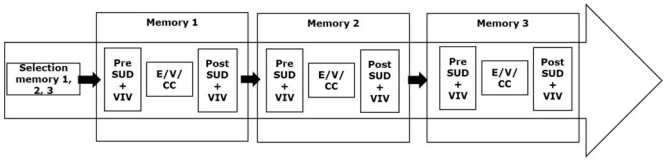

Fig. 1. The experiment timeline for Experiment 1 ($N = 30$). (Abbreviations: SUD, SUD scores; VIV, vividness scores; E, EMDR; V, VSDT; CC, Control Condition).

Immediately after completing all three conditions the participants were given a debriefing, in which they received more information about the experiment, and were asked to not disclose any information to others. In the final stage, the students were given their financial reimbursement or student credits. The procedure from memory selection to the post rating is visually depicted in Fig. 1.

2.1.4. Design

The experiment had a 3 (Condition: Eye Movement Desensitization and Reprocessing [EMDR] and Visual Schema Displacement [V] and Control Condition [CC]) by 2 (Time: pre-and post intervention) repeated measures design. Dependent variables were "Subjective Units of Disturbance" (SUD), and "Vividness" of the most disturbing part of the memory. To eliminate order effects and to control for carry-over effects, the order of the conditions was counterbalanced. There was also a counterbalance in order of the selected memories on disturbance, i.e. there was an alternation in the order based on the extent to which the memories felt disturbing. To control for experimenter effects, all conditions were carried out by the same research assistant.

2.1.5. Data analyses

A power analysis showed that to obtain a power of 0.8 with a probability error of 0.05 and an estimated effect size of 0.25, 36 participants would be needed. The sample size was calculated using G*power(ANOVA, repeated measures, within factors, 3 groups, 2 repetitions and a 0.5 correlation among measures). The data were analyzed using both Bayesian statistics and classical statistics.

2.1.5.1. Bayesian Statistics. Different expectations about the means were formulated as informative hypotheses (provided below) that were evaluated using Bayesian model selection. An advantage of the Bayesian approach is that it does not depend on dichotomous decisions (significant Y/N), but provides relative support for a pre-specified model or models (Klugkist, Laudy, & Hoijtink, 2005). This enables direct testing of theoretical expectations without the need for *post hoc* pairwise comparisons. Furthermore, different models can be tested at once. The results of the Bayesian model selection are expressed in terms of Bayes factors (BFs). A BF represents the level of evidence for one model compared to another. To evaluate the support for an informative hypothesis, it can be evaluated against a model without constraints. A BF lower than 1 indicates no support, whereas BFs above 1 indicate support for the informative hypothesis obtained from the data. A higher BF represents more support for the hypothesis. A BF of 1 means that both models receive equal support, and therefore that the data are indecisive with respect to the informative hypothesis. Finally, one can compare two informative hypotheses mutually by computing the ratio of the two BFs that evaluated the informative hypotheses against the unconstrained model. The Bayesian approach was added to the classical one, not because the authors believed it would favor any particular hypothesis, but because of the inherent problems in H0 testing that are reduced by a Bayesian approach (Mulder, Hoijtink & Klugkist, 2010; van de Schoot, Hoijtink, & Doosje, 2009; van de Schoot et al., 2014). Analyses were performed using the software BIEMS (see Mulder, Hoijtink, & de Leeuw, 2012; Mulder et al., 2010). Wetzels et al.

(2011) and Krypotos, Blanken, Arnaudova, Matzke, and Beckers (2017) provide more background information on Bayesian analyses and the comparison with p-value significance testing.

In the present experiment four models were compared for both SUD and vividness-scores. All analyses were conducted separately on SUD and Vividness. In all models, the assumption was made that the EMDR condition would outperform the CC. This expectation is based on earlier laboratory studies and clinical studies which showed the added value of EM in reprocessing emotional disturbing memories (e.g., Lee & Cuijpers, 2013). As this was the first VSDT experiment, no theoretical assumptions on the effect could be made. Nevertheless, there was no assumption that VSDT would have a detrimental outcome compared to the CC condition. The four compared models were:

1: (VSDT pre-post) > (EMDR pre-post) > (CC pre-post)
2: (VSDT pre-post) = (EMDR pre-post) > (CC pre-post)
3: (EMDR pre-post) > (VSDT pre-post) > (CC pre-post)
4: (EMDR pre-post) > (VSDT pre-post) = (CC pre-post)

2.1.5.2. Classical analyses. Besides Bayesian statistics, data were analyzed using SPSS (Version 22.0). A repeated measures (3 × 2) ANOVA was conducted, with Condition (VSDT; EMDR; CC) and Time (pre-and post) as independent variables. The dependent variables were SUD and vividness scores. All analyses were conducted separately on SUD and Vividness. *Post hoc* analyses were conducted when significant main or interaction effects were found.

2.2. Results

Data were obtained from 38 participants, although data from 8 participants were excluded, resulting in including 30 participants in the analyses. One student reported psychotic memories and was paranoid during the experiment (more specifically, the participant thought the research assistant had recorded the whole procedure). The other student was excluded because he stated he had been meditating during the CC. Six other patients were excluded because of procedural reasons. These participants reached “SUD 0” – meaning there was no emotionality left – in the VSDT condition before the 8 min were over (varying from between 7 min and 20 s to 7 min and 55 s). This resulted in a procedural error where the participants ended with the question posed at the end of the VSDT procedure. This meant the wording differed slightly from the wording of the question for those who participated in the full 8 min. Excluding these participants from the analyses meant leaving out participants who responded the quickest to the VSDT condition which, if anything, would be detrimental to VSDT compared to the other conditions.

Participants had the VSDT procedure carried out 2 to 4 times (the ‘Whoosh!’ was carried out 3 times per procedure; $M = 2.77$, $SD = 0.63$) and sets of EMDR (performed 30 s) 5 to 8 times ($M = 6.60$, $SD = 0.86$). Table 1 displays the means and standard deviations of the pre- and post-SUD and vividness scores.

2.2.1. Emotionality scores

2.2.1.1. Bayesian analyses. Results of the Bayesian analyses are shown in Table 2. These showed a clear pattern in that VSDT was superior to the EMDR condition and the CC, whereas the EMDR condition was found to be superior to the CC. This was supported by the BF value for Model 1 ($BF_1 = 6.54$). No support was found for superiority of the EMDR condition over both VSDT and the CC. Furthermore, no support was found for Model 2.

2.2.1.2. Classical analyses. A oneway repeated measures ANOVA showed that SUD pre-scores did not differ between the three conditions (F (2, 58) = 0.38, $p = .69$). A 2 × 3 repeated measures ANOVA showed there was a main effect for Time, F (1, 29) = 91.33, $p < .001$, $\eta_p^2 = 0.76$ and Condition, F (2, 58) = 10.37, $p < .001$, $\eta_p^2 = 0.26$ and a crucial interaction effect of Time x Condition, F (2, 58) = 29.99, $p < .001$, $\eta_p^2 = 0.51$. *Post hoc* pairwise comparisons adjusted with Bonferroni correction indicated that VSDT difference scores outperformed the EMDR condition ($p < .001$, $d = 0.84$) and the CC ($p < .001$, $d = 1.7$), and that the EMDR condition outperformed the CC ($p = .02$, $d = 0.69$).

2.2.2. Vividness scores

2.2.2.1. Bayesian analyses. BF's are summarized in Table 2. There was support for Model 2 and Model 3, which indicated superiority of the EMDR condition and VSDT over CC, with only slightly more evidence for the model showing equal efficacy of the EMDR condition and VSDT (BF_2: $BF_3 = 1.17$). There was also some support for Model 1, but no support for Model 4.

2.2.2.2. Classical analyses. Pre-vividness scores did not differ per condition F (2, 58) = 1.41, $p = .25$. For vividness, the results showed a main effect for Time F (1, 29) = 63.96, $p < .001$, $\eta_p^2 = 0.69$, Condition F (2, 58) = 8.91, $p < .001$, $\eta_p^2 = 0.24$ and the interaction effect of Time x Condition F (2, 58) = 7.61, $p < .001$, $\eta_p^2 = 0.21$ was also significant. *Post hoc* comparisons found the VSDT and EMDR conditions to be equally effective in reducing vividness of the emotional memories ($p = 1.00$), whereas VSDT and EMDR both outperformed the CC ($p = .01$, $d = 0.82$, $p = .01$, $d = 0.89$ respectively).

2.3. Discussion of experiment 1 and introduction to experiment 2

The results of Experiment 1 showed that the VSDT and EMDR conditions were superior to the CC in reducing emotionality and vividness. Surprisingly, considering the lack of rationale of VSDT and the fact that it was compared to an abbreviated version of an effective protocol (EMDR), VSDT was just as effective as the EMDR condition in reducing vividness and more effective than the EMDR condition in reducing emotionality. The pattern was statistically reflected in both Bayesian and classical analyses. As mentioned earlier, in the VSDT condition 6 out of 36 participants reported a SUD score of 0 before the 8 min that ended the intervention. Although for these early responders the question about the SUD score had been formulated as “How much is the disturbance reduced? It was ‘X’, what is it now?” it should have been formulated as “When you think about the worst part, how unpleasant does it feel or how much disturbance do you feel on a scale from 0, no disturbance, to 10, maximal disturbance?”. This is because this was the final question for all participants in all three conditions. Note also that only one of the participants reached “SUD 0” in the EMDR condition and none reached “SUD 0” in the CC. The participant with the “SUD 0” score in the EMDR condition was one of the six that was removed. Removal of the 6 subjects who responded so quickly and favorably to VSDT has, arguably, induced an underestimation of the VSDT effect. To determine the robustness of the results, a sensitivity analysis was conducted by which we compared the results with and without the rapid responders. Although larger SUD decreases were found when the rapid responders were included in the analyses, no differences in effect were found. To avoid any discussion about the integrity of the findings we decided to remove these individuals from the data set.

Table 1
Means and standard deviations pre- and post-measurement scores of SUD and vividness per condition ($N = 30$).

Condition	Pre-SUD	Post-SUD	Pre-vividness	Post-vividness
EMDR	7.40 (1.33)	5.10 (1.83)	7.97 (1.54)	6.07 (1.60)
VSDT	7.50 (1.28)	3.50 (1.93)	7.63 (1.54)	5.97 (1.87)
CC	7.27 (1.17)	6.20 (1.81)	8.07 (1.08)	7.47 (1.48)

Table 2
Bayes Factors (BF's) for emotionality and vividness of the tested models.

Model	Emotionality	Vividness
1. VSDT (pre-post) > EMDR (pre-post) > CC (pre-post)	6.54	1.92
2. VSDT (pre-post) = EMDR (pre-post) > CC (pre-post)	0.01	4.53
3. EMDR (pre-post) > VSDT (pre-post) > CC (pre-post)	0.00	3.86
4. EMDR (pre-post) > VSDT (pre-post) = CC (pre-post)	0.00	0.14

To test the robustness of the observations we decided to carry out a replication experiment. The question arose whether, if replicated, the favorable effect of VSDT would survive the passage of time. In order to evaluate the long-term effects, the design of Experiment 2 was changed into a between-group design. To be fully transparent about data collection, hypotheses and analyses, the replication experiment was pre-registered on the Open Science Framework (https://osf.io/kuenp/).

3. Experiment 2

3.1. Methods

3.1.1. Participants

Data from 75 participants (53 female; $M = 22.93$ years, $SD = 3.03$) were obtained and included in the analyses. They joined the experiment in exchange for course credits or financial compensation. Initially, 77 participants entered, but 2 participants were excluded from the analyses. One did not finish the procedure and one applied mindfulness during the control condition, which was considered a violation of the study protocol.

3.2. Materials

The second experiment also used SUD scores and vividness scores (see "Materials" Experiment 1 for a detailed description). The laughter point scores were not considered in this experiment, because the scores were not to be used in the analyses.

3.3. Procedure

Study procedures were approved by the Facultairy Medical Ethics Committee of the Utrecht University (The Netherlands) (FETC17-030). Procedures for Experiment 2 were identical to Experiment 1, except that participants were exposed to one randomly assigned intervention rather than three. The allocation to the conditions was sequential. Exposure to one intervention resulted in a shorter time span of approximately 30 min. After the checklist, one emotional memory was selected which was rated on a SUD scale from 0 ("not disturbing") to 10 ("maximum disturbance"). Memories rated with a SUD score of ≥7 were included in the experiment. In contrast to Experiment 1 "being a student" was not part of the inclusion criteria. The experiment was conducted by a graduate student who was not part of the author team and who was trained in VSDT by the VSDT developers, and in EMDR by an EMDR Europe accredited trainer. The student conducted all conditions. Fidelity checks were carried out on a video-recorded pilot sample to ensure the procedure was carried out properly.

After finishing the experiment, participants were asked for contact details and a follow-up (FU) telephone appointment was planned in a time range of 6–8 days after the intervention. The research student phoned the participants, and asked the participants to re-rate the SUD and vividness of the worst part of their memory. Next, a debriefing was send to them by email, which contained more information about the experiment. Participants were asked not to disclose any information to other persons, and the financial reimbursement or student credits were accredited. The procedure from memory selection to the post rating is visually depicted in Fig. 2.

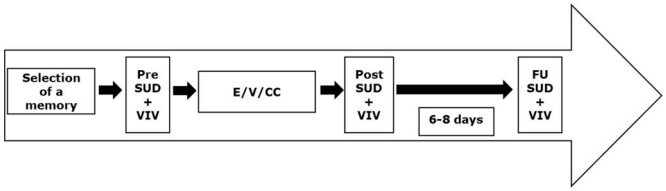

Fig. 2. The procedure timeline for Experiment 2 ($N = 75$). (Abbreviations: SUD, SUD scores; VIV, vividness scores; E, EMDR; V, VSDT; CC, Control Condition; FU, Follow-Up.)

3.3.1. Design

The study had a 3 (Condition: abbreviated Eye Movement Desensitization and Reprocessing (EMDR), Visual Schema Displacement (V) and Control Condition (CC) by 3 (time; pre, post and follow-up) between-repeated measures design, with Time as the within-subjects factor and Condition as the between-subjects factor. Dependent variables were "Subjective Units of Disturbance" (SUD; Emotionality) and "Vividness" of the memory.

3.4. Data analyses

A power analysis showed that to obtain a power of 0.8 with a probability error of 0.05 and an estimated effect size of 0.3, 75 participants would be needed. The sample size was calculated with G*power and the setting for ANOVA, repeated measures, between factors was used, with 3 groups and 3 repetitions and a 0.5 correlation among measures. The effect size was set at 0.3, which was slightly higher than in Experiment 1, due to the results found in Experiment 1.

3.4.1. Bayesian Statistics

Similar to Experiment 1 the data were analyzed using both Bayesian statistics and classical statistics. Because of the outcome of Experiment 1, two models were added (Models 5 and 6), with the outcome of the EMDR condition being equal to the outcome of the CC. All models were used for comparing both SUD and Vividness from pre-to post, from pre-to follow-up and from post to follow-up. Although no difference was expected from post to follow-up – which would only require Model 6 - for completeness also Model 1–5 were tested.

1: VSDT > EMDR > CC
2: VSDT = EMDR > CC
3: EMDR > VSDT > CC
4: EMDR > VSDT = CC
5: VSDT > EMDR = CC
6: EMDR = VSDT = CC

3.4.2. Classical analyses

The data were analyzed using SPSS (Version 22.0). The repeated measures -between subjects (3 × 3) ANOVA was conducted, with Condition (VSDT; EMDR; CC) and Time (pre; post and follow-up) as independent variables. The dependent variables were SUD and Vividness scores. Post hoc analyses were conducted when significant main or interaction effects were found.

Table 3
Means and standard deviations of the pre, post and follow-up measurement scores of SUD and vividness per condition.

Condition	Pre-SUD	Post-SUD	FU-SUD	Pre-vividness	Post- vividness	FU-Vividness
EMDR	7.92 (.70)	5.76 (1.42)	5.48 (1.45)	8.68 (1.18)	6.92 (1.35)	5.88 (1.74)
VSDT	7.84 (.80)	3.92 (1.98)	4.60 (1.29)	8.08 (1.26)	5.92 (2.10)	5.44 (1.47)
CC	7.64 (1.00)	6.56 (1.16)	6.00 (1.56)	8.00 (1.50)	7.28 (1.57)	6.72 (1.57)

Table 4
Bayes Factors (BF's) for emotionality and vividness of the tested models.

Model	Emotionality			Vividness		
	Pre- Post	Pre- FU	Post- FU	Pre- Post	Pre- FU	Post- FU
1: VSDT > EMDR > CC	6.41	4.85	0.01	5.17	2.27	0.57
2: VSDT = EMDR > CC	0.13	1.05	0.04	3.96	4.36	1.82
3: EMDR > VSDT > CC	0.03	0.33	0.04	1.40	3.92	1.85
4: EMDR > VSDT = CC	0.00	0.02	0.19	0.10	0.22	3.64
5: VSDT > EMDR = CC	0.76	1.26	0.03	0.72	0.10	0.88
6: EMDR = VSDT = CC	0.00	0.05	0.22	0.13	0.06	2.98

3.5. Results

Data were obtained from 75 participants. Participants had the VSDT procedure carried out 2 to 5 times (the 'Whoosh!' was carried out 3 times per procedure; $M = 3.28$, $SD = 0.89$) and sets of EMDR (performed 30 s) 5 to 8 times ($M = 6.16$, $SD = 0.84$). Table 3 displays the means and standard deviations of the pre, post and follow-up SUD and Vividness scores.

3.5.1. Emotionality scores

3.5.1.1. Bayesian analyses. Bayesian analyses, as displayed in Table 4, show the same pattern for the results from pre-to post as in Experiment 1. VSDT was superior to both the EMDR condition and the CC, while the EMDR condition was superior to the CC. This was supported by the BF values for Model 1. There was no support for superiority of the EMDR condition over both VSDT and the CC. There was also no support for equality of the VSDT and EMDR condition over the CC. In addition, the supplementary models in Experiment 2 were not supported by the results. When comparing the pre-SUD scores with the follow-up SUD most support was found for Model 1: VSDT > EMDR > CC. Although the support for Models 2 and 5 was no longer below 1, Model 1 was still the model that received most positive support. No support was found either for all other models. From post to follow-up the data did not support any of the models, indicating that there was another pattern or model, not pre-specified in this study, which could support the data better.

3.5.1.2. Classical analyses. A oneway ANOVA revealed that the SUD pre-measurement scores did not significantly differ per condition (F (2, 72) = 0.74, $p = .48$). A 3 (Time) by 3 (Condition) ANOVA was conducted revealing a main effect for Time, (F (2, 144) = 111.14, $p < .001$, $\eta_p^2 = 0.61$), and Condition (F (2, 72) = 12.79, $p < .001$, $\eta_p^2 = 0.26$), and a significant Time x Condition effect (F (4, 144) = 9.91, $p < .001$, $\eta_p^2 = 0.22$). A oneway ANOVA on difference scores from pre-to post showed an effect for Condition (F (2, 72) = 17.70, $p < .001$, $\eta_p^2 = 0.33$). *Post hoc* analyses adjusted with Bonferroni correction showed that VSDT outperformed both the EMDR condition ($p = .001$, $d = 0.93$) and the CC ($p < .001$, $d = 1.52$), but no difference was found between the EMDR condition and the CC ($p = .08$). From pre-to follow-up a oneway ANOVA on difference scores showed an effect for Condition (F (2, 72) = 6.60, $p = .002$, $\eta_p^2 = 0.16$). *Post hoc* analyses adjusted with Bonferroni correction showed a significant difference between VSDT and the CC ($p < .01$, $d = 0.97$), which was in favor of VSDT, but no significant difference was found between the VSDT and EMDR conditions ($p = .22$). In addition, no difference was found between the EMDR condition and the CC ($p = .22$). To analyze the effect on the SUD from post to follow-up a oneway ANOVA on the SUD difference scores from post to follow-up was carried out showing a significant effect for Condition (F (2, 72) = 4.17, $p = .02$, $\eta_p^2 = 0.10$). The *post hoc* analyses adjusted with Bonferroni correction showed a difference between VSDT and the CC ($p = .02$, $d = 0.70$), with a significant better result for the CC. No difference was found between the EMDR condition with either the VSDT condition and the CC ($p = .11$, $p = 1.00$).

3.5.2. Vividness scores

3.5.2.1. Bayesian analyses. See Table 4 for the results of the Bayesian analyses. Bayesian analyses of pre-to post scores showed the strongest support for Model 1, reflecting the fact that the VSDT condition outperformed the EMDR condition and that the EMDR condition was superior to the CC. There was also support for Model 2 (VSDT = EMDR > CC), but there was a relative benefit for Model 1 of (BF1: BF2) 1.30. There was less, but still some support for Model 3 (EMDR > VSDT > CC). For Models 4, 5 and 6 no support was found. When comparing the pre-vividness scores with the vividness scores at follow-up most support was found for Model 2: VSDT = EMDR > CC, implying that the VSDT and EMDR condition were equally effective and both more effective than the CC in decreasing vividness from pre-to FU. But there was also support for Model 3 (EMDR > VSDT > CC). Model 2 had only slightly more relative support than Model 3 (BF2: BF3 = 1.11). Model 1 (VSDT > EMDR > CC) also had a BF which indicated support. Similarly, in Model 1, 2 and 3 both the EMDR and VSDT condition were superior to the CC. No support was found for Model 4, 5 and 6. From post to the follow-up measurement the results corroborated Model 4 (EMDR > VSDT = CC) the most, but there was also support for Model 6 (VSDT = EMDR = CC), with only slightly more evidence for the EMDR condition being more successful in the decrease in vividness while VSDT and the CC were equal in their effect (BF3: BF6 = 1.22). There was also some support for Model and Model 2. The data did not corroborate Model 1 and Model 5.

3.5.2.2. Classical analyses. A oneway ANOVA showed pre-vividness scores did not differ per condition (F (2, 72) = 1.99, $p = .15$). A 3 (Time) by 3 (Condition) ANOVA was conducted and revealed there was a main effect for Time (F (2, 144) = 65.64, $p < .001$, $\eta_p^2 = 0.48$), Condition (F (2, 72) = 3.64, $p = .03$, $\eta_p^2 = 0.09$), and the interaction effect of Time x Condition (F (4, 144) = 3.72, $p = .01$, $\eta_p^2 = 0.09$) was also significant. A oneway ANOVA on Vividness difference scores from pre-to post showed an effect for Condition (F (1, 72) = 5.15, $p = .01$, $\eta_p^2 = 0.13$). Post hoc analyses adjusted with Bonferroni correction

revealed no differences between the conditions VSDT and EMDR ($p = 1.00$) and between EMDR and the CC ($p = .83$), but VSDT and the CC differed significantly ($p = .01$, $d = 0.82$). A oneway ANOVA on Vividness difference scores from pre-to FU showed an effect of Condition (F (2, 72) = 6.38, $p = .003$, $\eta_p^2 = 0.15$). *Post hoc* analyses adjusted with Bonferroni correction showed that both the VSDT and EMDR condition were superior to the CC ($p = .01$, $d = 0.84$, $p \leq .01$, $d = 0.54$) but that the EMDR and VSDT condition did not differ ($p = 1.00$). A oneway ANOVA on Vividness difference scores from post to the follow-up measurement showed no effect for Condition (F (2, 72) = 0.64, $p = .53$).

3.6. Discussion

The purpose of the present study was to critically test the claim that VSDT is capable of reducing the emotionality and vividness of negative memories in healthy participants. As well as comparing VSDT to a non-active control condition, we also determined whether the effects would be similar, or even superior, to those of an abbreviated version of the EMDR protocol. Two experiments were conducted. In the first experiment VSDT was compared to an abbreviated version of EMDR and a Control condition (CC). In the second (replication) experiment a follow up measurement was added to determine whether the results would be maintained over time. Both experiments showed VSDT was superior to the EMDR condition and the CC, while the EMDR condition had stronger effects than the CC in reducing emotional disturbance of disturbing memories. VSDT and the EMDR condition were equally effective in reducing vividness and both were shown to be more effective than the CC.

The results were unexpected, robust and in favor of VSDT. In terms of emotionality scores, the only time-span where VSDT did not excel, was in effecting SUD scores from post to the follow-up measurement in the second experiment. This raises the question as to why VSDT was so effective. The procedure is curious, to say the least, whereas the originators of this method have little clue about how the procedure might work. We aimed to determine the differences between VSDT with an active control condition (an abbreviated version of EMDR) with a paradigm often used in experimental psychopathology research. However, there are a few hypotheses which might partially explain the effectiveness of VSDT. First, the unconventional nature of VSDT may have generated, at the very beginning of this treatment, the *expectation* that this would be helpful. But it is far from obvious that unconventionality breeds positive expectations. Also, empirically it was shown in an experimental set-up like the present one, that participants with (experimentally induced) negative expectations about recall + EM reported the same positive effects as individuals with induced positive expectations (Littel, van Schie, & van den Hout, 2017b). Second, much of the effects of EMDR can conveniently be explained by the working memory theory (e.g. Gunter & Bodner, 2008; van den Hout & Engelhard, 2012). For reasons of parsimony it is tempting to try and explain VSDT effects in the same theoretical terms. On the one hand, one could argue that this effort to do so is farfetched, stretching working memory theory beyond its limits. On the other, VSDT has a number of procedural steps ensuring that while the memory is activated, patients' working memory is taxed. This is particularly the case during the interval when the participants eyes are directed at a point of which he/she indicates that the emotions attached to the memory are most strongly felt, and the 'Whoosh!' is introduced. The latter step has a number of similarities with the desensitization phase of EMDR therapy, but both procedures seem to differ strongly in terms of intensity. That is, whereas in EMDR, after the memory has been activated, the therapist's hand is introduced to help employ horizontal eye movements to tax working memory resources, the VSDT therapist applies a massive dose of working memory taxation by using a combination of an unannounced exclamation of the word 'Whoosh!' (i.e., an auditory task), and an unexpected and quick oblique movement of his arm (i.e., visual task). Third, there is another unique aspect of VSDT that might at least partially explain some of its effectiveness, and that is the addition of arousal. Unlike EMDR, in VSDT distress and arousal is created since the patient does not know when the therapist will suddenly call out loud the word 'Whoosh!' and he has to quickly follow the therapist's arm movement. To this end, the latter is particularly interesting given that a recent study showed that individuals who first underwent a stress inducing task to increase their arousal showed reduced vividness of emotionally neutral memories after the dual tasking, while those who carried out a recall only task, did not (Littel, Remijn, Tinga, Engelhard, & van den Hout, 2017a). In the same vein, a placebo controlled study showed that reducing arousal by beta blockers reduces the positive effects of Recall + EM (Littel, Kenemans, et al., 2017). In future studies it would be enlightening to test several variations of the procedure to tease out which of the various elements (e.g. blinking, eye closure, saying "Whoosh") contributes to the effects. Testing variations lends itself to dissection of (non)-effective ingredients.

Some limitations need to be noted. Firstly, the follow-up period of the second experiment was rather short (6–8 days). It is unclear what would have happened to the SUD scores if the follow-up period had been longer, or if the telephone interview had been replaced by a face-to-face follow-up interview. Although a face-to-face interview may have been more reliable, this was not undertaken due to the anticipated drop out. However, it is unlikely that this would explain why the EMDR condition and the CC show better results than VSDT from post to the follow-up measurement, as the same type of assessment was made in these conditions. Secondly, the measurements taken consisted of self-reports about subjective mental states, i.e., ratings of experienced vividness and emotionality. Note however that these very same measures are well accepted in memory research (Benjamin et al., 2010; Bremner et al., 1999; Devilly & Spence, 1999; Heeren, Reese, McNally, & Philippot, 2012; Osuch et al., 2001; Littel, Remijn, et al., 2017) and the SUD scale shows good psychometric properties to measure emotionality of memories (Kim, Bae, & Chon Park, 2008). Therefore there is no reason to doubt the usefulness of such measures in the experiments reported here. Yet, it should be noted that our sample was not a clinical sample, and that the participants were not screened for psychopathology. Even more so, they were excluded from the study when they had a current psychiatric treatment. Although the same pattern in healthy participants and patients fulfilling the diagnostic criteria of PTSD is observed in EMDR research, the results of this study are not generalizable to individuals fulfilling the diagnostic criteria of PTSD or any other mental health condition without reservation. Furthermore, we did not apply the full standard protocol. Because of time restraints we used an abbreviated lab protocol which was based upon the exact text of the EMDR Standard protocol.

A strength of the study is the design of the experiments, which was rather strict and comparable to experiments carried out in related areas (Engelhard, van den Hout, Janssen, & van der Beek, 2010; Engelhard, van den Hout, & Smeets, 2011; van Veen, Engelhard, & van den Hout, 2016; van Veen et al., 2015). It is difficult to see how the methods used may have favored the positive outcomes for VSDT. In fact, and if anything, the method may have been detrimental towards the VSDT-condition. For example, the exclusion of the six most positive responders to the VSDT condition from the analyses in the first experiment, could easily have led to an underestimation of the true VSDT effects. Furthermore, applying a strict Bonferroni correction for multiple testing is a conservative strategy (Rothman, 2014) and the fact that VSDT survived most of the Bonferroni corrected comparisons, argues *a fortiori* for the effectiveness of the intervention. Moreover, the VSDT condition was at a disadvantage compared to the EMDR condition, as a larger part of the time of the intervention was used for preparation and explanation. The duration of strict VSDT was thereby somewhat smaller than the duration of the desensitization in the EMDR condition. This translated in a lower number of VSDT sets ($M = 2.77$, $M = 3.28$, respectively) compared to EMDR sets ($M = 6.60$, $M = 6.16$, respectively), averaged

over both experiments. This is of a lesser disadvantage when duration of the conditions would have been longer or preparations and general instructions preceded the start of the conditions. On the other hand, apparently 8 min were enough to find a statistically significant effect. Furthermore, in case of EMDR, a wide array of experimental studies has shown that even very short exposures to parts of the EMDR procedure (e.g., 96 s) is effective in reducing emotionality (e.g., Gunter & Bodner, 2008). Regarding VSDT the present study was the first that investigated the efficiency of the procedure which means that no fixed format existed prior to our study. That is why a study protocol for the experiments was developed in close collaboration with the originators of VSDT. In future research the duration of the intervention certainly needs to be taken into account. Conceivably, the optimal dose varies with certain patient characteristics, level of disturbance or type of memory and is therefore an important issue to address. During both the VSDT and EMDR condition participants were regularly asked to report SUD scores related to the target, but the formulations, derived from the two protocols, were slightly different. In the VSDT condition participants were asked to what extent the emotional response to the target (i.e. SUD-score) was reduced compared to the last reported rating, while in the EMDR condition the participants were asked to determine the SUD score in relation to the image representing the worst part of memory. This difference in procedural aspects of both interventions could – just like any other part of the procedures – have contributed to reporting a decrease and be channeling in a certain direction. However, these process instructions are part of the procedures and could, therefore, not be removed from the protocols. This was the exact reason why we removed the fast responders from the analysis in Study 1. It should be noted that at pre-test, post-test and follow-up, instructions and questions were identical in all conditions. If the formulation of this instruction in the VSDT condition resulted in an effect in that condition, this would mean that this is an important ingredient of the procedure and working mechanism. Further research should clarify whether this is the case or not. Another strength was the use of two forms of statistical analyses, which contributes to the robustness of our findings. In the first study, there was a great overlap between the Bayesian and Classical analyses, and this was also true for the second study, albeit that small differences were found. Because of the relative support that Bayesian Statistics can provide we tend to favor the conclusion based upon the Bayesian Statistics.

In summary then, in two laboratory experiments, we studied effects of an unusual procedure that was held to reduce the emotionality and vividness of unpleasant memories in healthy participants. Had we not found any effects we may have been reluctant to report the data to the scientific community: nothing new with falsifying yet another power-therapy's claim. Effects were, however, observed: They were powerful and survived replication. A further step should be to test the procedure in patients to see whether VSDT is also capable in affecting aversive memories in patients. Also, measurements of PTSD symptoms should be included as well as longer follow-up periods to determine whether the effects are long-lasting. It may be too early to call for a large RCT as yet, although the latter would be warranted if controlled n = 1 studies and/or well-documented clinical case series were to corroborate the present data.

Conflict of interest and funding

Ad de Jongh receives income from published books on EMDR therapy and for training postdoctoral professionals in this method. None of the other authors have a conflict of interest in this study. A small grant from the Dutch EMDR association was awarded in May 2017 to Suzy Matthijssen. Irene Klugkist is supported with a Vidi grant (452-12-010) from the Netherlands Organization for Scientific Research (NWO). Marcel van den Hout is supported by a TOP grant (number: 40-00812-98-12030) from the Netherlands Organization for Health Research and Development (ZonMw).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbtep.2018.11.006.

References

American Psychiatric Association (2004). *Practice guideline for the treatment of patients with acute stress disorder and posttraumatic stress disorder.* Arlington, VA: American Psychiatric Association Practice Guidelines.

Andrade, J., Kavanagh, D., & Baddeley, A. (1997). Eye-movements and visual imagery: A working memory approach to the treatment of post-traumatic stress disorder. *British Journal of Clinical Psychology, 36*(2), 209–223. https://doi.org/10.1111/j.2044-8260.1997.tb01408.x.

Benjamin, C. L., O'Neil, K. A., Crawley, S. A., Beidas, R. S., Coles, M., & Kendall, P. C. (2010). Patterns and predictors of subjective units of distress in anxious youth. *Behavioural and Cognitive Psychotherapy, 38*(4), 497–504. https://doi.org/10.1017/S1352465810000287.

Bisson, J. I., Ehlers, A., Matthews, R., Pilling, S., Richards, D., & Turner, S. (2007). Psychological treatments for chronic post-traumatic stress disorder. *The British Journal of Psychiatry, 190*(2), 97–104. https://doi.org/10.1192/bjp.bp.106.021402.

Bleich, A., Kotler, M., Kutz, I., & Shalev, A. (2002). *Guidelines for the assessment and professional intervention with terror victims in the hospital and in the community.* Jerusalem, Israel: National Council for Mental Health.

Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry, 162*(2), 214–227.

Bremner, J. D., Narayan, M., Staib, L. H., Southwick, S. M., McGlashan, T., & Charney, D. S. (1999). Neural correlates of memories of childhood sexual abuse in women with and without posttraumatic stress disorder. *American Journal of Psychiatry, 156*(11), 1787–1795.

Chen, Y. R., Hung, K. W., Tsai, J. C., Chu, H., Chung, M. H., Chen, S. R., ... Chou, K. R. (2014). Efficacy of eye-movement desensitization and reprocessing for patients with posttraumatic-stress disorder: A meta-analysis of randomized controlled trials. *PLoS One, 9*(8), e103676. https://doi.org/10.1371/journal.pone.0103676.

CREST (2003). *The management of post traumatic stress disorder in adults. Belfast, Northern Ireland: Clinical resource efficiency team of the Northern Ireland department of health.* Social Service and Public Safety.

Cuijpers, P., & Cristea, I. A. (2016). How to prove that your therapy is effective, even when it is not: A guideline. *Epidemiology and Psychiatric Sciences, 25*, 428–435.

De Jongh, A., Ernst, R., Marques, L., & Hornsveld, H. (2013). The impact of eye movements and tones on disturbing memories of patients with PTSD and other mental disorders. *Journal of Behavior Therapy and Experimental Psychiatry, 44*, 447–483.

De Jongh, A., & ten Broeke, E. (2012). *Handboek EMDR: Een geprotocolleerde behandelmethode voor de gevolgen van psychotrauma (Handbook of EMDR: A protocolized treatment method for the consequences of psychotrauma).* Amsterdam: Pearson Assessment and Information B.V.

Devilly, G. J., & Spence, S. H. (1999). The relative efficacy and treatment distress of EMDR and a cognitive-behavior trauma treatment protocol in the amelioration of posttraumatic stress disorder. *Journal of Anxiety Disorders, 13*(1), 131–157. https://doi.org/10.1016/S0887-6185(98)00044-9.

Dutch National Steering Committee Guidelines Mental Health (2003). *Multidisciplinary guideline anxiety disorders.* Utrecht, the Netherlands: Quality Institute Health Care CBO/Trimbos Institute.

Engelhard, I. M., van den Hout, M. A., Janssen, W. C., & van der Beek, J. (2010). Eye movements reduce vividness and emotionality of "flashforwards". *Behaviour Research and Therapy, 48*(5), 442–447.

Engelhard, I. M., van den Hout, M. A., & Smeets, M. A. (2011). Taxing working memory reduces vividness and emotional intensity of images about the Queen's Day tragedy. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(1), 32–37. https://doi.org/10.1016/j.jbtep.2010.09.004.

Gunter, R. W., & Bodner, G. E. (2008). How eye movements affect unpleasant memories: Support for a working-memory account. *Behaviour Research and Therapy, 46*(8), 913–931.

Haute, Autorité de la Santé (2007). *Guidelines for the treatment of long term psychiatric conditions – serious anxiety disorders.* Paris, France: HAS.

Heeren, A., Reese, H. E., McNally, R. J., & Philippot, P. (2012). Attention training toward and away from threat in social phobia: Effects on subjective, behavioral, and physiological measures of anxiety. *Behaviour Research and Therapy, 50*(1), 30–39. https://doi.org/10.1016/j.brat.2011.10.005.

Herbert, J. D., Lilienfeld, S. O., Lohr, J. M., Montgomery, R. W., T O'Donohue, W., Rosen, G. M., et al. (2000). Science and pseudoscience in the development of eye movement desensitization and reprocessing: Implications for clinical psychology. *Clinical Psychology Review, 20*(8), 945–971. https://osf.io/kuenp/.

Kim, D., Bae, H., & Chon Park, Y. (2008). Validity of the subjective units of disturbance scale in EMDR. *Journal of EMDR Practice and Research, 2*(1), 57–62. https://doi.org/10.1891/1933-3196.2.1.57.

Klugkist, I., Laudy, O., & Hoijtink, H. (2005). Inequality constrained analysis of variance: A Bayesian approach. *Psychological Methods, 10*(4), 477–493.

Krypotos, A. M., Blanken, T. F., Arnaudova, I., Matzke, D., & Beckers, T. (2017). A primer on Bayesian analysis for experimental psychopathologists. *Journal of Experimental Psychopathology, 8*(2), 140. https://doi.org/10.5127/jep.057316.

Lee, C. W., & Cuijpers, P. (2013). A meta-analysis of the contribution of eye movements in processing emotional memories. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(2), 231–239. https://doi.org/10.1016/j.jbtep.2012.11.001.

Lileinfeld, S. O. (1996). EMDR treatment: Less than meets the eye. *The Skeptical Inquirer, 20*, 25–31.

Littel, M., Kenemans, J. L., Baas, J. M., Logemann, H. A., Rijken, N., Remijn, M., ... van den Hout, M. A. (2017c). The effects of β-adrenergic blockade on the degrading effects of eye movements on negative autobiographical memories. *Biological Psychiatry, 82*, 8. https://doi.org/10.1016/j.biopsych.2017.03.012.

Littel, M., Remijn, M., Tinga, A. M., Engelhard, I. M., & van den Hout, M. A. (2017a). Stress enhances the memory-degrading effects of eye movements on emotionally neutral memories. *Clinical Psychological Science, 5*(2), 316–324. https://doi.org/10.1177/2167702616687292.

Littel, M., van Schie, K., & van den Hout, M. A. (2017b). Exploring expectation effects in EMDR: Does prior treatment knowledge affect the degrading effects of eye movements on memories? *European Journal of Psychotraumatology, 8*(sup1), 1328954. https://doi.org/10.1080/20008198.2017.1328954.

Lohr, J. M., Kleinknecht, R. A., Tolin, D. F., & Barrett, R. H. (1995). The empirical status of the clinical application of eye movement desensitization and reprocessing. *Journal of Behavior Therapy and Experimental Psychiatry, 26*(4), 285–302.

Maxfield, L., Melnyk, W. T., & Hayman, G. C. (2008). A working memory explanation for the effects of eye movements in EMDR. *Journal of EMDR Practice and Research, 2*(4), 247–261. https://doi.org/10.1891/1933-3196.2.4.247.

Meyers, M. A. (2007). *Happy accidents: Serendipity in modern medical breakthroughs.* Arcade Publishing.

Mulder, J., Hoijtink, H., & de Leeuw, C. (2012). BIEMS: A Fortran 90 program for calculating Bayes factors for inequality and equality constrained models. *Journal of Statistical Software, 46*(2), 1–39.

Mulder, J., Hoijtink, H., & Klugkist, I. (2010). Equality and inequality constrained multivariate linear models: Objective model selection using constrained posterior priors. *Journal of Statistical Planning and Inference, 140*(4), 887–906. https://doi.org/10.1016/j.jspi.2009.09.022.

National Institute for Clinical Excellence (2005). *Post traumatic stress disorder (PTSD): The management of adults and children in primary and secondary care.* London, UK: NICE Guidelines.

Osuch, E. A., Benson, B., Geraci, M., Podell, D., Herscovitch, P., McCann, U. D., et al. (2001). Regional cerebral blood flow correlated with flashback intensity in patients with posttraumatic stress disorder. *Biological Psychiatry, 50*(4), 246–253. https://doi.org/10.1016/S0006-3223(01)01107-6.

Rothman, K. J. (2014). Six persistent research misconceptions. *Journal of General Internal Medicine, 29*(7), 1060–1064. https://doi.org/10.1007/s11606-013-2755-z.

van de Schoot, R., Hoijtink, H., & Doosje, S. (2009). Rechtstreeks verwachtingenevalueren of de nul hypothese toetsen? nul hypothese toetsing versus bayesiaanse model selectie [Directly evaluating expectations or testing the null hypothesis: Null hypothesis testing versus Bayesian model selection]. *De Psycholoog, 4*, 196–203. https://doi.org/10.2139/ssrn.1919023.

van de Schoot, R., Kaplan, D., Denissen, J. J. A., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. G. (2014). A gentle introduction into Bayesian analysis: Applications to developmental research. *Child Development, 85*(3), 842–860. https://doi.org/10.1111/cdev.12169.

Seidler, G. H., & Wagner, F. E. (2006). Comparing the efficacy of EMDR and trauma focused cognitive behavioral therapy in the treatment of PTSD: A meta-analytic study. *Psychological Medicine, 36*, 1515–1522. https://doi.org/10.1017/S0033291706007963.

Shapiro, F. (1989). Eye movements desensitization. A new treatment for post-traumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 20*, 211–217.

Shapiro, F. (1995). *Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures.* New York: Guilford Press.

van den Hout, M. A., & Engelhard, I. M. (2012). How does EMDR work? *Journal of Experimental Psychopathology, 3*(5), 724–738. https://doi.org/10.5127/jep.028212.

van Veen, S. C., Engelhard, I. M., & van den Hout, M. A. (2016). The effects of eye movements on emotional memories: Using an objective measure of cognitive load. *European Journal of Psychotraumatology, 7*(1), 30122. https://doi.org/10.3402/ejpt.v7.30122.

van Veen, S. C., van Schie, K., Wijngaards-de Meij, L. D., Littel, M., Engelhard, I. M., & van den Hout, M. A. (2015). Speed matters: Relationship between speed of eye movements and modification of aversive autobiographical memories. *Frontiers in Psychiatry, 6*. https://doi.org/10.3389/fpsyt.2015.00045.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E. J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science, 6*(3), 291–298. https://doi.org/10.1177/1745691611406923.

Wilson, S. A., Becker, L. A., & Tinker, R. H. (1995). Eye movement desensitization and reprocessing (EMDR) treatment for psychologically traumatized individuals. *Journal of Consulting and Clinical Psychology, 63*(6), 928–937. https://doi.org/10.1037/0022-006X.63.6.928.

Wolpe, J. (1969). *The practice of behavior therapy* (2nd ed.). New York: Pergamon Press.

World Health Organization (2013). *Guidelines for the management of conditions that are specifically related to stress.* Geneva, Switzerland: WHO.