Behaviour Research and Therapy 92 (2017) 94–105

Contents lists available at ScienceDirect

Behaviour Research and Therapy

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

Eye movement during recall reduces objective memory performance: An extended replication

Arne Leer [a, *], Iris M. Engelhard [a], Bert Lenaert [b, c, d], Dieter Struyf [b], Bram Vervliet [b, e, f], Dirk Hermans [b]

[a] Clinical Psychology, Utrecht University, Heidelberglaan 1, 3584 CS, Utrecht, The Netherlands
[b] Centre for the Psychology of Learning and Experimental Psychopathology, KU Leuven - University of Leuven, Leuven, Belgium
[c] School for Mental Health and Neuroscience, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands
[d] Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[e] Department of Psychiatry, Harvard Medical School, Boston, MA, USA
[f] Department of Psychiatry, Massachusetts General Hospital, Boston, MA, USA

ARTICLE INFO

Article history:
Received 12 September 2016
Received in revised form
28 February 2017
Accepted 6 March 2017
Available online 9 March 2017

Keywords:
EMDR
PTSD
Eye movements
Memory performance
Fear generalization

ABSTRACT

Eye Movement Desensitization and Reprocessing (EMDR) therapy for posttraumatic stress disorder involves making eye movements (EMs) during recall of a traumatic image. Experimental studies have shown that the dual task decreases self-reported memory vividness and emotionality. However valuable, these data are prone to demand effects and little can be inferred about the mechanism(s) underlying the observed effects. The current research aimed to fill this lacuna by providing two objective tests of memory performance. Experiment I involved a stimulus discrimination task. Findings were that EM during stimulus recall not only reduces self-reported memory vividness, but also slows down reaction time in a task that requires participants to discriminate the stimulus from perceptually similar stimuli. Experiment II involved a fear conditioning paradigm. It was shown that EM during recall of a threatening stimulus intensifies fearful responding to a perceptually similar yet non-threat-related stimulus, as evidenced by increases in danger expectancies and skin conductance responses. The latter result was not corroborated by startle EMG data. Together, the findings suggest that the EM manipulation renders stimulus attributes less accessible for future recall.

Eye movement desensitization and reprocessing (EMDR) is a treatment that successfully reduces symptoms of posttraumatic stress disorder (PTSD; $d = 1.43$, 95% CI [1.02, 1.83], Bradley, Greene, Russ, Dutra, & Westen, 2005), outperforms wait-list control conditions ($d = 1.25$, 95% CI [-0.97, 3.48], Bradley et al., 2005), and is equally effective as trauma-focused cognitive-behavioral therapy (TFCBT; Bisson et al., 2007; Bradley et al., 2005; Seidler & Wagner, 2006). Hence, it is recommended as a treatment-of-choice for PTSD (e.g., APA, 2004; NICE, 2005). However, about one-third of PTSD patients does not show clinical improvement (Bradley et al., 2005) and little is known about EMDR's mechanisms of change (e.g., Gunter & Bodner, 2008; Leer, Engelhard, & van den Hout, 2014). Gaining a better understanding of the processes that account for EMDR's outcomes is essential for treatment optimization in terms of efficacy, efficiency, patient selection, and individualization of treatment (Kazdin, 2007).

For a long time, many scholars were skeptical about the introduction of EMDR therapy, and particularly the vague theoretical rationale and lack of empirical support for the eye movement component (see Engelhard, 2012; e.g., Herbert et al., 2000; McNally, 1999; Lilienfeld, 1996; Lohr, Kleinknecht, & Tolin, & Barrett, 1995). Early reviews suggested that eye movements were not an essential component of treatment and that EMDR may be effective because it contains an exposure component. Recent years have seen a steep increase in experimental research addressing these issues (for a review, see Van den Hout & Engelhard, 2012). The laboratory model for investigating the EM component comprises three phases (Van den Hout, Muris, Salemink, & Kindt, 2001). At pre-test, healthy participants select and briefly recall a negative autobiographical memory, and rate its vividness and emotionality. During a subsequent intervention phase they visualize the memory for a fixed period of time, either with or without EM. Finally, at post-test, they

* Corresponding author.
E-mail address: A.Leer@uu.nl (A. Leer).

http://dx.doi.org/10.1016/j.brat.2017.03.002
0005-7967/

retrieve and re-rate the memory. A recent meta-analysis suggests that EM benefits – in terms of reductions in memory vividness/emotionality – are large in such experimental trials ($d = 0.74$, 95% CI [0.57, 0.91]), and small to medium in clinical trials comparing EMDR treatment with EM to EMDR treatment without EM ($d = 0.27–0.41$, 95% CI [0.07–0.13, 0.47–0.70]; Lee & Cuijpers, 2013).

Several explanations of these findings have been put forward (e.g., Gunter & Bodner, 2008). For example, the *investigatory-reflex* account purports that EM induces a strong sense of relaxation that can last up to 10 min and becomes associated with the trauma memory (e.g., Kuiken, Bears, Miall, & Smith, 2002). EM, however, was shown to be beneficial when conducted concurrently, but not just before memory recall (Gunter & Bodner, 2008), which contradicts the theory. Alternatively, the *increased hemispheric communication* account explains that EM during memory recall facilitates communication between the left and right brain hemispheres, which enhances memory retrieval and desensitization (Christman, Garvey Propper, & Phaneuf, 2003). However, vertical EM was demonstrated to be as effective as horizontal EM (Gunter & Bodner, 2008). Importantly, these findings are accommodated by a *working memory* account (Andrade, Kavanagh, & Baddeley, 1997). This theory posits that EM taxes working memory and thus competes for limited resources that are demanded by memory recall. As a result, recall of a memory while making EMs is less vivid and evokes less extreme emotional responses. Notably, changes in memory phenomenology are not only observed *during* EM (Andrade et al., 1997; exp. 4.; Kavanagh, Freese, Andrade, & May, 2001), but also immediately afterwards, i.e. when the cognitive load caused by EM had been removed, and at 24-h (Leer et al., 2014) and 1-week follow-up (Gunter & Bodner, 2008; exp. 2). EM benefits thus extend beyond the experimental/clinical session and corroborate the clinical observation of trauma memory amelioration following EMDR.

Robust and promising as the EM effects are, it has been acknowledged that most prior studies have relied on self-report measures, which are prone to demand effects (e.g., Kearns & Engelhard, 2015). Demand effects may also be expected in the control condition that involves mere recall (without EM), because most participants in the relevant studies are psychology students who are generally familiar with the clinical efficacy of imaginal exposure. Yet, most studies have not found strong effects of mere recall. Also, Gunter and Bodner (2008) demonstrated EM benefits when EM were performed during recall, but not when EM were performed before recall, which challenges a 'demand characteristics account'. Nevertheless, collecting non-self-report data, such as physiological or behavioral data, is essential for two reasons. First, it enables an evaluation of the scientific integrity of the existing and to be collected self-report data (Van den Hout, Bartelski, & Engelhard, 2013). Already, several changes in (Dutch) EMDR guidelines for treatment and training of therapists have been applied on the basis of self-report data (Beer et al., 2011). Cross-validation of previous research findings is thus valuable both theoretically and clinically. Second, non-self-report data may advance our understanding of *how* the EM component adds to EMDR's effectiveness. It has been proposed that EMs exert their long-term effects because they cause a loss in memory detail through memory reconsolidation (e.g., Maxfield, Melnyk, & Hayman, 2008; Van den Hout et al., 2010). This hypothesis, however, cannot be addressed by the mere assessment of subjective memory ratings, but rather calls for objective indexes of memory performance.

Non-self-report data have been reported in a small number of studies. One experiment demonstrated pre- to post-test reductions in potentiation of the startle blink reflex during negative ideation in an EM condition, but not in a control condition (Engelhard, van Uijen, & van den Hout, 2010). A second experiment in individuals with performance anxiety showed that EM during imagery of generic fear-related scripts, but not such imagery without EM, caused pre- to post-test heart rate deceleration (Kearns & Engelhard, 2015). Two studies including physiological endpoints did not collect data during separate pre- and post-tests involving memory recall. Barrowcliff, Gray, Freeman, and MacCulloch (2004) reported decreases in skin conductance in the course of the EM intervention. Schubert, Lee, and Drummond (2011) reported reduced heart rate from pre- to post-intervention baseline periods that did not comprise concurrent memory recall. However valuable, such experimental designs hamper the interpretation of changes in memory.

In addition, so far two studies have collected objective behavioral data (Van den Hout et al., 2013; Van Schie, Engelhard, & van den Hout, 2015). In the study by Van den Hout et al. (2013), participants encoded two detailed drawings of scenes. Subsequently, half of the participants recalled one of the images with EM, the other half recalled one of the images without EM. All participants were then presented with cutouts of the encoded images and with cutouts of other, unseen images. They were asked to indicate as fast and as accurate as possible whether each cutout was part of one of the encoded images. Results were that making EM during image recall, but not recall only, increased response latencies. It seemed unlikely that the result represented a shift in the speed-accuracy trade-off from speed to accuracy, because there was no effect of EM on decision accuracy. If response latencies reflect the extent to which memory attributes are retrievable, this experiment shows that the effect of EM on memory performance is not confined to self-report data. However, given that the images were hedonically positive (see Van den Hout, Eidhof, Verboom, Littel, & Engelhard, 2014), it is unclear whether the findings can be generalized to EMDR. In an extended replication by van Schie et al. (2015), participants learned relationships between neutral words and aversive pictures. Next, via a cued-recall procedure they were presented a cue word and prompted to recall the associated image, with or without EM. Finally, cue words were presented on screen surrounded by four cutouts of images. Participants were asked to identify which cutout had previously been paired with each cue word. The intervention failed to produce pre- to post-test changes in memory vividness and emotionality, and did not affect response latencies in the matching task. Presumably, suboptimal memory retrieval during the intervention explains their null finding.

Overall, it may be concluded that non-self-report data on EM are needed, yet scarce. The goal of the present study was to fill this lacuna. First, we set out to provide a conceptual replication of Van den Hout et al. (2013). To this end, we used a novel discrimination task (Experiment I). The second aim was to extend the findings by testing whether EM during recall reduces the objective performance of a hedonically negative memory. To this end, we used a fear generalization paradigm (Experiment II).

1. Experiment I

Experiment I examined the effect of EM on memory performance using a stimulus discrimination task. Following an encoding phase, participants recalled the image of a neutral male face either with EM (experimental condition) or without (control condition). Before and after the intervention they rated memory vividness. Directly after the intervention they were presented pictures of novel faces that perceptually resembled the original face, and indicated whether or not these images were identical to the original one. Based on Van den Hout et al. (2013) we predicted that EM during recall, relative to recall only, increases response latencies.

Fig. 1. Target images form sets A and B were recalled in Experiment I. Test images from sets A and B were presented in the stimulus discrimination task of Experiment I. Target image and test image 2 from set A were used as CS and GS, respectively, in Experiment II.

2. Method

2.1. Participants

Participants were 27 volunteers (26 undergraduates), who received remuneration or course credits. The experiment was approved by the local ethical committee. A priori exclusion criteria were pregnancy, serious medical conditions, (past or present) psychiatric diagnoses, having an electronic implanted device (e.g., a pacemaker), and pain or problems related to hands or wrists. After the experiment, knowledge about EMDR was assessed: one participant had linked the experimental procedure to the EM employed in EMDR, and was excluded. The final sample comprised 26 participants (age: $M = 19.96$, $SD = 2.96$; 5 males).

2.2. Stimulus material

Two different image sets ('A' and 'B') were created by selecting images of neutral male faces (511 × 768 pixels) from the Radboud Faces Database (Langner et al., 2010). Each set comprised one 'target image' and three 'test images' (see Fig. 1). Test images were morphs between the target image and other neutral male faces (created with Abrosoft Fantamorph software) and resembled the target image in decreasing steps of perceptual similarity.

2.3. Stimulus discrimination task

The stimulus discrimination task comprised six trials during which one of three test images – from either set A or B – was presented for 8 s. Image order was semi-random, starting with three different images. Inter-trial-intervals ranged between 15 and 25 s ($M = 20$ s). Instructions were that several images of faces would be presented and that one should indicate as quickly and correctly as possibly whether the face at hand was identical to the image they had seen and recalled in previous phases of the experiment. Response options were pressing a green key ('same') or a red key ('different') on a keyboard. Participants could press only once and each image remained on the monitor for exactly 8 s.

2.4. Procedure

Participants were seated in a dimly lit room about 42 cm in front of a 17-inch monitor (1440 × 900 pixels). They received oral information about the study and provided written informed consent. The order of the experimental and control conditions was counterbalanced within subjects. When image set A was used in the experimental condition, then image set B was used in the control condition, and vice versa, which was balanced as well. Each condition consisted of five phases (see Fig. 2). During phase one, participants were presented the target image for 30 s and instructed to encode as much detail as possible, because they would be asked some questions about the image later on. During phase two, they brought the target image to mind as vividly as possible and rated its mental representation by filling out a 100 mm visual analogue scale (VAS: 0 = *not vivid at all*; 100 = *very vivid*). During phase three, they recalled the target image for 24 s[1] while watching a black screen (control condition) or while making EM by following a white dot (Ø = 1 cm) that moved horizontally across a black screen at one cycle per second (experimental condition). In the experimental condition, and only during this phase, the distance to the monitor was reduced to 30 cm to ensure a substantial eye movement angle (approximately 53°). After the recall period participants in both conditions were instructed to focus on something else for 10 s. This was done to minimize the risk that, at the post vividness test, participants, would refer to how they experienced the image during the intervention phase. During phase four, they retrieved the target image and re-rated its vividness. Phase five involved the stimulus discrimination task. At the end of the experiment, participants were debriefed, asked about their

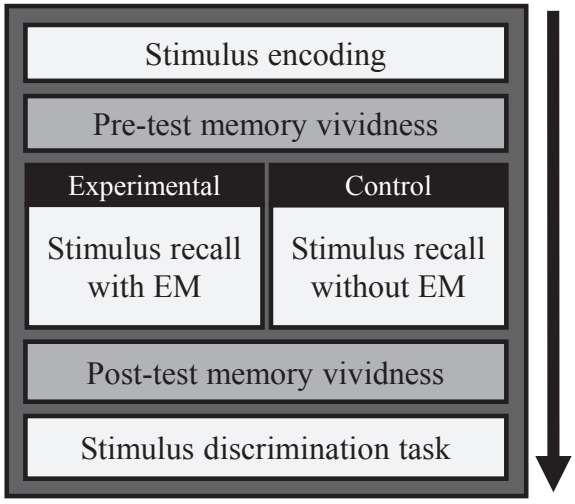

Fig. 2. Experimental procedure of Experiment I.

[1] Typically, experimental studies that aim to desensitize negative autobiographical memories employ four or more 24-s sets of eye movements (e.g., Gunter & Bodner, 2008; Van den Hout et al., 2001; note that sets of such duration are rooted in the EMDR protocol; Shapiro, 2001). Because the current experiment employed de novo stimulus material that was not yet consolidated in long-term memory at the time of the intervention, and based on two pilot studies, we decided on using only one 24-s set.

Table 1
Mean raw reaction times (SDs) in milliseconds.

	Test image 1	Test image 2	Test image 3
Experimental	2895.60 (1689.69)	3202.25 (1822.27)	2486.96 (1620.74)
Control	2701.60 (1454.59)	2898.04 (1803.68)	2138.62 (1282.86)

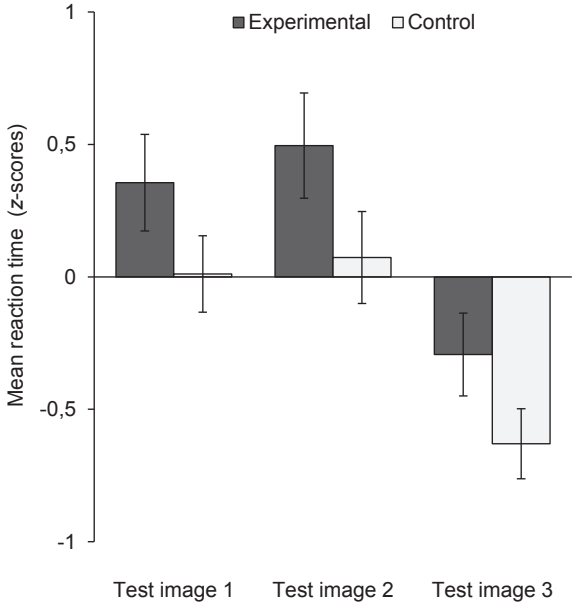

Fig. 3. Mean standardized reaction times. Error bars reflect standard errors of the mean.

knowledge on EMDR, and compensated.

3. Results

3.1. Memory vividness

A 2 × 2 repeated measures ANOVA was conducted with Condition and Time as within-subjects variables. There were main effects of Time, $F(1, 25) = 51.46$, $p < 0.001$, $\eta p2 = 0.673$, and Condition, $F(1, 25) = 12.21$, $p = 0.002$, $\eta p2 = 0.328$. Crucially, there was a Time × Condition interaction, $F(1, 25) = 7.92$, $p = 0.009$, $\eta p2 = 0.243$. In line with our expectations, reductions in memory vividness following EM during recall (pre-test: $M = 78.00$, $SD = 11.19$; post-test: $M = 63.38$, $SD = 14.24$) were stronger as compared to mere recall (pre-test: $M = 79.46$, $SD = 10.47$; post-test: $M = 71.69$, $SD = 12.56$).

3.2. Discrimination speed

Table 1 shows the raw reaction time (RT) data collected in the discrimination task. To account for inter-individual differences in reaction speed, standardized RTs (z-scores) were analyzed (see Fig. 3). A 2 × 2 repeated measures ANOVA was conducted with Condition and Image number as within-subjects variables. The crucial main effect of Condition was significant, $F(1, 25) = 12.06$, $p = 0.002$, $\eta p2 = 0.325$. Consistent with the first hypothesis, overall responses were slower in the experimental condition. There was also a main effect of Image number, $F(2, 50) = 14.31$, $p < 0.001$, $\eta p2 = 0.364$, with participants classifying test image 3 faster compared to test images 1 and 2. This result reflects that participants found test image 3 easiest to classify, which agrees with the fact that test image 3, relative to test images 1 and 2, shares less features with the target image (see Fig. 1), and corresponds with the finding that nearly all participants correctly classified test image 3 as dissimilar from the target image (see Table 2). No interaction effect was found, $F < 1$.

3.3. Discrimination accuracy

Table 2 shows the categorization responses collected in the discrimination task. As each test image was shown twice, participants could give two responses per test image. Thus, they could give two 'same as target' responses, two 'different from target' responses, or one 'same as target' and one 'different from target' response.

Recall that test image 1 was perceptually closest to the target image, followed by test image 2, and that test image 3 had the least resemblance (see Fig. 1). Table 2 suggests that, overall, accuracy in discriminating test images from the target image increased as a function of perceptual dissimilarity between the two image types. However, Fisher-Freeman-Halton exact tests revealed that frequencies per category did not significantly differ by group[2] (test image A1, $p = 0.053$; the other test images, $ps > 0.199$). Thus, there was no statistical evidence that the EM procedure affected discrimination accuracy.

One interpretation of the same/different category is that it reflects participants' uncertainty as to whether the image at hand was or was not similar to the target image. Accordingly, we explored if the EM procedure resulted in participants becoming more uncertain. For test image A1, a majority of the experimental group (64%) and a minority of the control group (17%) fell in the same/different category. Fisher's exact test revealed that this difference was significant, $p = 0.036$ (all other test images, $ps > 0.378$).

4. Discussion of Experiment I and introduction to Experiment II

In sum, findings in Experiment I were that EM during recall, relative to recall only, reduces memory vividness and memory performance during a subsequent stimulus discrimination task, as evidenced by reduced response latencies, but not by reduced response accuracy. These results provide a first conceptual replication of Van den Hout et al. (2013). Similar to van den Hout and colleagues, we found that the dual task slows down the process of stimulus discrimination, while leaving discrimination accuracy intact. This pattern of results rules out an explanation in terms of a shift in the speed–accuracy trade–off (i.e., from being as fast as possible to becoming more accurate). Together, these studies demonstrate that the effects of EM on memory are not confined to self-report and preclude a demand characteristics account of the widely reported EM effects (cf. Engelhard, van Uijen et al., 2010; Kearns & Engelhard, 2015; Van den Hout et al., 2013). They further extend previous investigations by showing that the dual task not only affects autobiographical memories, but also mental representations of relatively simple, non-idiosyncratic stimulus material. This is especially relevant for future examinations of the EM effects that wish to employ a (fear) conditioning paradigm and/or require the presentation of simple and standardized stimuli. Two earlier investigations in which participants recalled de novo images failed to provide corroborating evidence (Leer, Engelhard, Dibbets,

[2] A within-subjects comparison of the two interventions would violate the chi-square assumption of independence. Therefore, between-group comparisons were made using the Freeman-Halton extension of the Fisher exact probability test, for each stimulus separately.

Table 2
Frequencies of categorization responses per test image and per intervention.

Responses	A1		A2		A3		B1		B2		B3	
	Exp	Ctrl	Exp	Ctrl	Exp	Ctrl	Exp	Ctrl	Exp	Ctrl	Exp	Ctrl
Same/same	3	4	1	2	0	0	5	8	2	2	1	0
Same/different	7	2	5	2	2	2	3	3	3	3	1	0
Different/different	1	6	7	9	10	11	3	2	5	6	9	13

& Van den Hout, 2013; Van Schie et al., 2015). Methodological differences may account for this inconsistency in findings. Leer et al. (2013) did not include an isolated encoding phase before the dual task and pre-test vividness scores were relatively low, leaving little room for improvement. Van Schie et al. (2015) tested the effects of cued recall rather than instructed recall, and their null results may be due to suboptimal memory retrieval. Together, these findings suggest that boundary conditions should be respected when including non-idiosyncratic stimulus material in the dual task paradigm.

A limitation of Experiment I was that hedonically neutral stimuli were used, which may limit the applicability of the findings to EMDR for traumatic memories. Therefore, in Experiment II we aimed to create hedonically negative stimuli. To this end, a fear conditioning paradigm was employed. Such paradigm is widely used for studying the acquisition and extinction of responses indicative of fear (e.g., Craske, Hermans, & Vansteenwegen, 2006). In a typical fear conditioning experiment the participant is presented with pairings of an initially neutral conditional stimulus (CS; e.g., a sound) with an intrinsically aversive unconditional stimulus (UCS; e.g., electrical stimulation). As a result of these pairings, the meaning of the CS changes, i.e. it becomes a threatening stimulus capable of evoking conditional fear responses (CRs; LeDoux, 2014). Such procedure thus creates hedonically aversive stimuli in a way that is controlled by the experimenter.

Clinical observations and analogue studies show that CRs are not confined to the CS involved in the original learning experience, but generalize to stimuli that in some way resemble the CS (for review, see: Dymond, Dunsmoor, Vervliet, Roche, & Vervliet, 2015). From a memory perspective, generalization can be considered a function of the (e.g., perceptual) resemblance between the stimulus at hand and the retrieved CS memory. More specifically, to the extent that CS attributes (e.g., shape or color) cannot be retrieved from memory, and therefore cannot be used to discriminate the CS from novel stimuli, fear generalization will occur (Riccio, Rabinowitz, & Axelrod, 1994; Struyf, Zaman, Vervliet, & Van Diest, 2015). Examples supporting this model are reduced stimulus discrimination following a retention interval, which suggests a causal role of forgetting of stimulus attributes (Bahrick, Clark, & Bahrick, 1967; Scapinello & Yarmey, 1970), and the finding that feature learning, i.e. paying extra attention to specific stimulus features during conditioning, affects what type of generalization stimuli evokes a CR (Vervliet, Kindt, Vansteenwegen, & Hermans, 2010; Vervliet & Geens, 2014).

The rationale in Experiment II was as follows. Confrontation with a stimulus that reminds one of a CS, thus signaling potential threat, will motivate the individual to fully activate and examine the CS memory, and to compare it with the stimulus at hand. Based on the findings in Experiment I and Van den Hout et al. (2013), it can be expected that following EM during CS recall this process will take more time. And, the longer it takes to classify an ambiguous stimulus as 'safe', the more likely it will be that people react as if it were 'dangerous' (i.e., adopt a better-safe-than-sorry strategy). Accordingly, we hypothesized that recall of a CS memory with EM, compared to without EM, increases the strength of fear generalization. This hypothesis is also relevant in light of current EMDR practices, as it contradicts what one would expect given that EMDR is a treatment for PTSD. Whereas the blurring of emotional 'hotspots' may have beneficial effects on stress-related complaints, loss in detail of memory for related (useful) information may be associated with increased fearful responding to objectively safe stimuli or situations. Results of Experiment II may thus be relevant to understanding EMDR treatment in more than one way.

First, participants underwent a fear conditioning procedure, in which a picture of a neutral male face (CS) was followed by an electrical shock (UCS), and participants rated shock expectancy. Next, they recalled the CS image with EM (experimental group) or without (control group), and rated memory vividness. Finally, during a test phase, they saw a novel face that perceptually resembled the original face (generalization stimulus; GS), and rated shock expectancy again. We predicted decreases in memory vividness and stronger generalization of fear from the CS to the GS in the experimental group. Fear was operationalized as self-reported shock-expectancy, elevation of electrodermal responding, and potentiation of the startle blink reflex.

5. Method

5.1. Participants

Participants were 54 volunteers (49 undergraduates), who received remuneration or course credits. The procedure was approved by the local ethical committee. Exclusion criteria were similar to Experiment I. The first participant was excluded because startle probes were not properly presented and only part of the data was saved. Another participant was excluded because of linking the experimental procedure to the EM employed in EMDR. The final sample comprised 52 participants (age: $M = 21.67$, $SD = 7.03$; 13 males). By order of appearance, they were allocated to an experimental ($n = 25$) or control group ($n = 27$).

5.2. Stimulus material

The standardized RT data of Experiment I were explored to see which test image was associated with the largest impact of the EM procedure. Test image 2 from image set A showed the largest difference in discrimination speed between the two interventions (experimental: $M = 0.42$, $SD = 0.32$; control: $M = -0.41$, $SD = 0.17$, $t(24) = 2.32$, $p = 0.032$, $d = 0.909$; the raw data indicated the same). Therefore, in Experiment II, the target image and test image 2 from image set A were used as CS and GS, respectively (see Fig. 1).

A 2-ms electrocutaneous stimulus was used as UCS. Shocks were generated via a constant current stimulator (Digitimer DS7A; Hertfordshire, UK) and delivered through a pair of AgCl electrodes (Coulbourn, 8-mm V91-01) that were filled with K-Y Jelly (Johnson & Johnson) and attached to the wrist of the dominant hand. Shock intensity was individually set (range: 2–64 mA; $M = 23.46$ mA, $SD = 15.11$; see *Procedure*).

5.3. Questionnaire

To control for inter-individual differences in baseline anxiety levels, which are known to affect fear learning (Grillon et al., 2006; Lissek et al., 2005), the State Trait Anxiety Inventory was administered (STAI-DY; Spielberger, Gorsuch, & Lushene, 1970). State anxiety and trait anxiety were both measured with 20 items that were scored on a 1-4-point scale (1 = *not at all*, 4 = *severely*).

5.4. Skin conductance responses

Skin conductance responses (SCRs) were recorded by attaching a pair of AgCl electrodes (Coulbourn, 8-mm V91-01) filled with K-Y Jelly (Johnson & Johnson) to the hypothenar palm of the non-preferred hand. The electrodes transmitted a constant voltage of 0.5 V that was applied and registered by an isolated skin conductance coupler (Coulbourn, V71-23). Data were sampled at 10 Hz.

5.5. Startle responses

Startle blink electromyographic (EMG) responses were elicited by binaural acoustic presentation of 50-ms white noise probes at 102 dB. EMG activity was recorded by placing two AgCl electrodes (Coulbourn, 4-mm V91-02) over the orbicularis oculi region of the right eye, and a third one (Coulbourn, 8-mm V91-01) approximately 3–4 cm superior to the upper borders of the inner brows (Fridlund & Cacioppo, 1986). These sites were first cleaned with exfoliating cream (Louis Widmer) and the electrodes were filled with microlyte electrode gel (Coulbourn, X11-71). The raw EMG signal was amplified by an isolated bioamplifier (Coulbourn, V75-04) with high pass (13 Hz) and low pass (1 KHz) filters. The signal was then rectified and smoothed by a multifunction integrator (Coulbourn, V76-24) with a time constant of 20 ms. Data were digitized at 1 KHz.

5.6. Procedure

Participants were seated in a dimly lighted room about 42 cm in front of a 17-inch monitor (1440 × 900 pixels). They were told that pictures of faces would appear on the monitor with breaks in between, that they would occasionally receive electrical shocks, and that it was their job to learn to predict shock occurrence. After providing written informed consent, they filled out the STAI. Next, electrodes were attached and shock intensity was individually set via a work up procedure (cf. Fonteyne, Vervliet, Hermans, Baeyens, & Vansteenwegen, 2009). The experimenter presented a series of shocks. Shock intensity started at 1 mA and was increased stepwise. Participants rated each shock on a scale ranging from 0 to 10 (0 = *I don't feel anything*; 1 = *I feel something, but this is not painful, it is just a sensation*; 2 = *it starts to feel painful, but it is still a very small pain*; 10 = *this is the maximum tolerable pain for me in this experiment*). Participants were urged to notify the experimenter when their maximum level was reached or when they wanted shock intensity to be turned down. They were further asked whether they would be able to tolerate occasional shocks at their chosen level. Lastly, headphones were put on. The experiment was programmed using Affect 4.0 software (Spruyt, Clarysse, Vansteenwegen, Baeyens, & Hermans, 2010) and comprised five consecutive phases: a practice, acquisition, intervention, test, and extinction phase (see Fig. 4).

5.6.1. Practice phase

There were four baseline trials and four image trials, which alternated, starting with a baseline trial. During each trial, a shock expectancy VAS (0 = *certainly no shock*; 100 = *certainly a shock*) appeared on screen, which disappeared after 6 s or 1 s after response registration. The CS was presented only during image trials for 6 s. Participants were told that this phase served to practice filling out the VAS and that no shocks would be delivered yet. They were encouraged to indicate shock expectancy as quickly as possible each time the VAS appeared on screen. Startle probes were presented 4 or 5 s after each VAS onset. This phase thus comprised a total of eight startle probe presentations that served as habituation trials. Inter-probe-intervals ranged between 17 and 25 s (21 s on average).

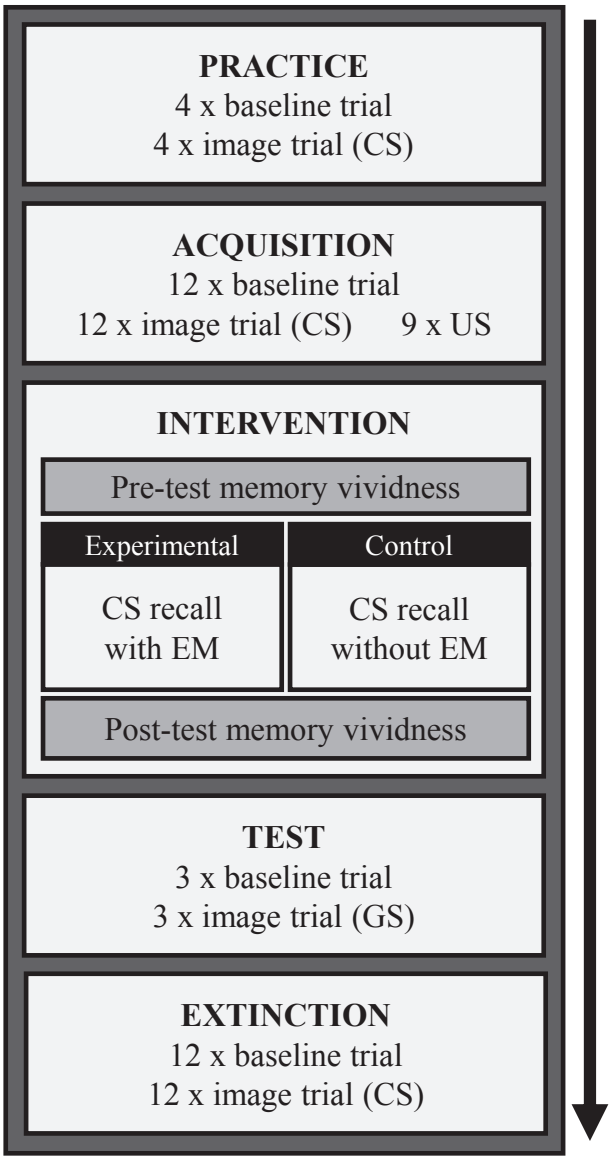

Fig. 4. Experimental procedure of Experiment II.

5.6.2. Acquisition phase

There were 12 baseline trials and 12 image trials that were similar to the practice trials. During this phase, nine randomly chosen image trials were immediately followed by shock (75% reinforcement schedule). Baseline trials were included to allow for examination of differential learning. Participants were told that they would occasionally receive electrical shocks during this phase and that it was their job to indicate shock expectancy each time the VAS appeared on screen.

5.6.3. Intervention phase

Shock electrodes were disconnected and participants were told that no shocks would be presented during this phase. They were then exposed to the CS for 10 s and received instructions to encode as much detail as possible because they would be asked some questions about the image later on. Following the same procedure as in Experiment I, participants then rated the vividness of the CS memory (pre-test), recalled it for 24 s with EM (experimental group) or without EM (control group), and finally provided a second vividness rating (post-test).

5.6.4. Test phase

Shock electrodes were reconnected, but no shocks were given during this phase. First, to habituate the startle response, four probes were presented with 17–25 s (M = 21 s) intervals in between while participants looked at a blank screen. Subsequently, participants were presented three baseline trials and three image trials, which alternated and started with a baseline trial. During the image trials, the GS appeared on screen for 6 s. Participants were told that this phase would start with a few startle probe trials, followed by presentations of the same face as in the previous phases and other, somewhat different, faces. They were further told that the original face would sometimes be followed by shock, but that the other faces would never be followed by shock.

5.6.5. Extinction phase

Without further instructions, the extinction phase started. This phase was similar to the acquisition phase, except that no shocks were delivered, and it served to invalidate the CS-UCS contingency. After the experiment, participants were debriefed and compensated for their time.

5.7. Data preparation

PSPHA software was used to analyze the physiological data (De Clercq, Verschuere, De Vlieger, & Crombez, 2006). SCRs were calculated by subtracting the baseline value (average SC level for the 2 s preceding stimulus onset) from the peak value that was recorded in the period of 0–4 s following VAS onset (Pineles, Orr, & Orr, 2009; note that startle probes occurred 4 s or 5 s after VAS onset). Minimum response amplitude was set at 0.02 μS. All other responses were scored as 0 and left in the analyses (Pineles et al., 2009). The data were T-transformed to adjust for inter-individual differences (50 + 10*[(raw score − M)/SD]; Dawson, Schell, & Filion, 2007; Haesen & Vervliet, 2015). Finally, block averages were calculated per three trials, resulting in four acquisition blocks and one test block.

Startle amplitudes were calculated by subtracting the baseline value (average EMG level 0–20 ms after probe onset) from the peak value that was identified in the period of 21–200 ms after probe onset. Each response was visually inspected. Trials that showed substantial noise during the baseline period (11.1%) or did not show a response to the probe (5.8%) were excluded from statistical analyses. To account for inter-individual differences, all data were T-transformed (Blumenthal et al., 2005). Finally, three-trial block averages were calculated, resulting in four acquisition blocks and one test block.

6. Results

Groups did not differ in state anxiety (experimental: M = 33.28, SD = 6.81; control: M = 31.67, SD = 9.75), t < 1, trait anxiety (experimental: M = 37.12, SD = 7.08; control: M = 34.30, SD = 7.72), $t(50) = 1.37$, $p = 0.176$, age, $t(50) = 1.39$, $p = 0.172$, gender ratio, $\chi^2(1, N = 52) = 0.642$, $p = 0.423$, or shock intensity (experimental: M = 22.92, SD = 12.76; control: M = 23.96, SD = 17.23), t < 1, indicating successful randomization.

6.1. Memory vividness

Repeated measures ANOVA showed a significant Time × Group interaction, $F(1, 50) = 5.55$, $p = 0.022$, $\eta p^2 = 0.100$. As expected, reductions over time were larger in the experimental condition (pre-test: M = 81.12, SD = 10.67; post-test: M = 60.76, SD = 17.14) than in the control condition (pre-test: M = 83.48, SD = 9.81; post-test: M = 71.70, SD = 14.61).

6.2. Shock expectancy

Recall that the fear acquisition phase was scheduled before the intervention phase, and that no group differences were expected in differential learning (see Fig. 5). This was investigated by a 2(Group) x 2(Stimulus: CS vs. baseline) x 12(Trial) repeated measures ANOVA. There were main effects of Stimulus, $F(1, 45) = 313.11$, $p < 0.001$, $\eta p^2 = 0.874$, and Trial, $F(11, 495) = 2.16$, $p = 0.040$, $\eta p^2 = 0.046$, that were qualified by a Stimulus × Trial interaction,

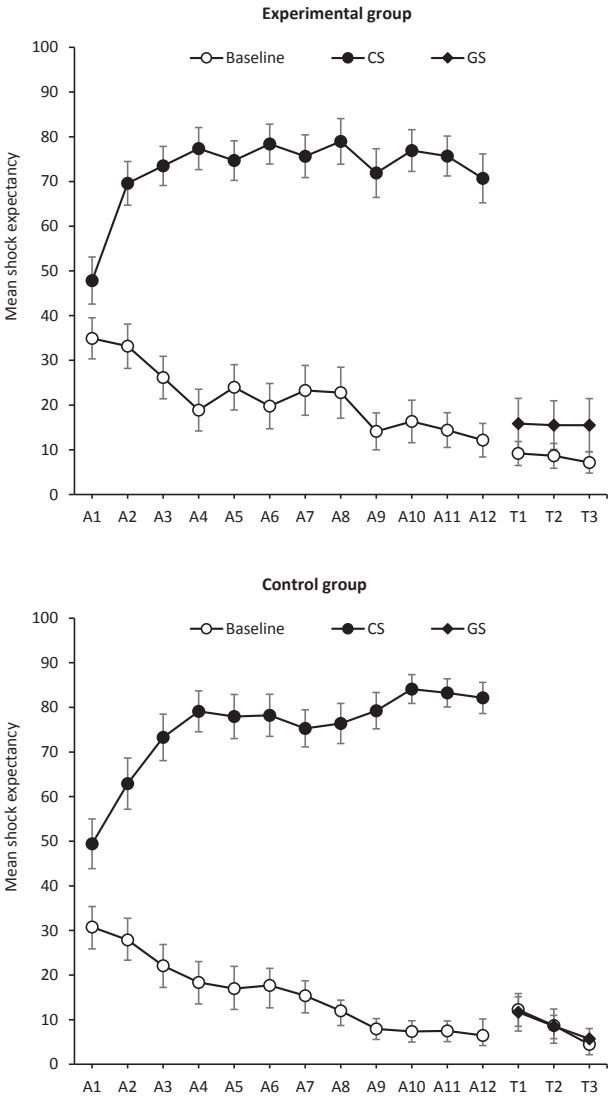

Fig. 5. Mean UCS expectancy ratings during the acquisition (A1-12) and test phase (T1-3). Error bars reflect standard errors of the mean.

$F(11, 495) = 14.10$, $p < 0.001$, $\eta p^2 = 0.239$. Differential shock expectancy increased over time. No other effects were found, $Fs < 1$.

To test the hypothesis we compared the two groups on strength of generalized shock expectancy, which was defined as expectancy during image test trials minus expectancy during baseline test trials. Fig. 5 suggests that at the end of the acquisition phase conditional shock expectancy was stronger in the control group compared to the experimental group, which was unanticipated. This observation was confirmed by a *t*-test that compared the two groups on the average difference score of the last three acquisition trials (experimental: $M = 59.70$, $SD = 33.49$; control: $M = 76.71$, $SD = 17.50$), $t(50) = 2.32$, $p = 0.029$, $d = 0.637$. Because the intensity of generalized responding is directly related to the intensity of conditional responding, this between-group difference hinders testing the hypothesis. Therefore, we corrected for this difference before analyzing the test phase data. For each participant, the difference scores of the last three acquisition trials and the three test trials were used to compute *z*-scores. These *z*-scores allow us to interpret the strength of (generalized) responding during the test phase as *relative* to the strength of (conditional) responding at the end of the acquisition phase. A 2(Group) x 3(Trial: test trials 1–3) repeated measures ANOVA on these standardized data only showed the expected main effect of Group, $F(1, 47) = 5.88$, $p = 0.019$, $\eta p^2 = 0.111$, other $Fs < 1$. This finding indicates that generalization of shock expectancy was stronger in the experimental group compared to the control group, and supports the hypothesis.

6.3. Skin conductance responding

Four participants did not show any response at all during the experiment and were excluded from analyses. First, differential learning was examined. A 2(Group) x 2(Stimulus) x 4(Block) repeated measures ANOVA on the acquisition phase data showed main effects of Group, $F(1, 46) = 5.40$, $p = 0.025$, $\eta p^2 = 0.105$, Stimulus, $F(1, 46) = 59.88$, $p < 0.001$, $\eta p^2 = 0.566$, and Block, $F(3, 138) = 40.78$, $p < 0.001$, $\eta p^2 = 0.470$. On average, SCR was higher in the control group, higher to CS trials than to baseline trials, which indicates successful learning, and habituated over time. No other effects were found, largest $F = 1.08$, $p = 0.358$.

Like for the shock ratings, we checked whether the strength of conditional responding during the last acquisition block was comparable between the two groups. There was no evidence for such a difference, $t < 1$. As a test of the hypothesis, we then compared the two groups on strength of generalized responding, i.e. the difference between SCR during the image test block and SCR during the baseline test block. Unexpectedly, no group difference was found, $t < 1$. Note, however, that Fig. 6 shows a sudden and substantial increase in SCR during the first baseline trial. This likely reflects an orienting response, as the test phase started with four noise alone trials and the very first presentation of the shock expectancy VAS occurred during the first baseline trial. Importantly, this effect hampers the detection of between-group differences when examining the three-trial block average. Fig. 6 further suggests that during test trial 2 generalized responding was stronger in the experimental group compared to the control group. This was confirmed by a *t*-test, $t(46) = 2.17$, $p = 0.035$, $d = 0.626$, providing further support for the hypothesis. This effect ceased at test trial 3, $t < 1$.

6.4. Startle responding

Differential learning was examined using a 2(Group) x 2(Stimulus) x 4(Block) repeated measures ANOVA. There were main effects of Stimulus, $F(1, 47) = 10.58$, $p = 0.002$, $\eta p^2 = 0.184$, and Block,

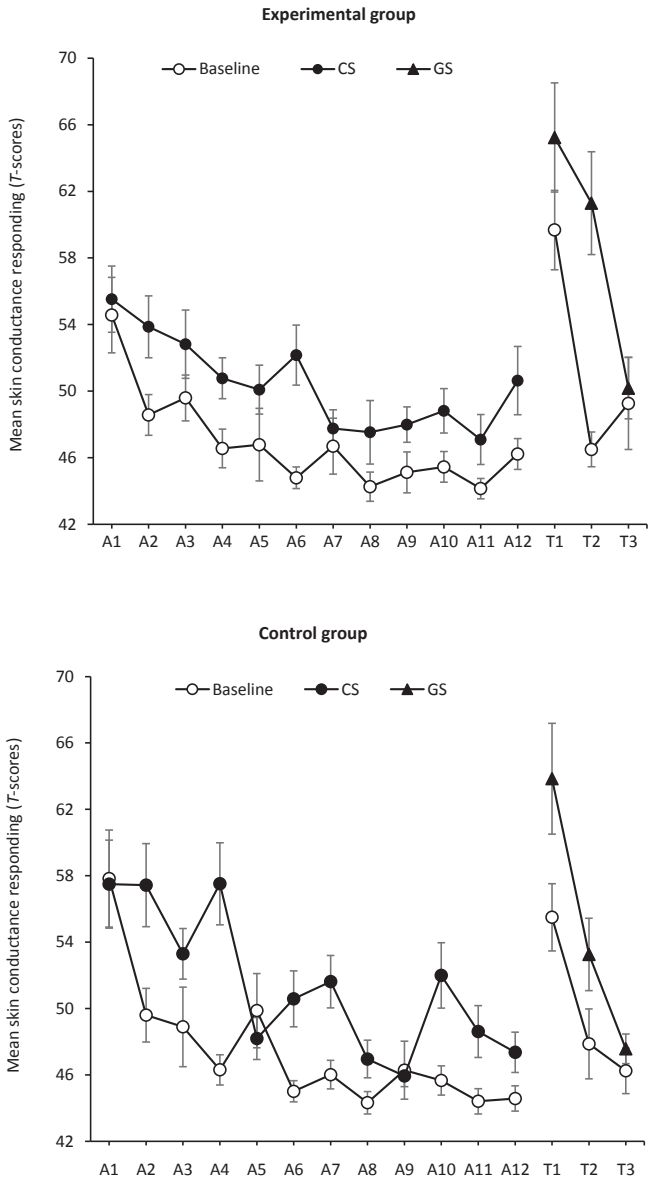

Fig. 6. Mean standardized skin conductance responses during the acquisition (A1-12) and test phase (T1-3). Error bars reflect standard errors of the mean.

$F(3, 141) = 52.58$, $p < 0.001$, $\eta p^2 = 0.528$, but not of Group, $F(1, 47) = 3.00$, $p = 0.090$. Startle responding was higher to CS trials than to baseline trials, indicating successful learning, and habituated over time. No other effects were found, largest $F = 1.17$, $p = 0.325$.

Before examining the hypothesis, we tested for a group difference in conditional responding during the last acquisition block, which was not found, $t < 1$. Next, to test the hypothesis, groups were compared on the strength of generalized responding, i.e. the difference between responding during the image test block and the baseline test block. Contrary to expectations, the intensity of differential responding was still similar between the groups during the test block, $t < 1$. Comparable to the SCR data, Fig. 7 suggests that the startle amplitude during the first baseline trial was affected by an orienting response, and that groups differed in generalized responding during the second test trial. The latter observation, however, was not confirmed, $t(34) = 1.53$, $p = 0.137$. Thus, the startle EMG data did not provide further support for the hypothesis.

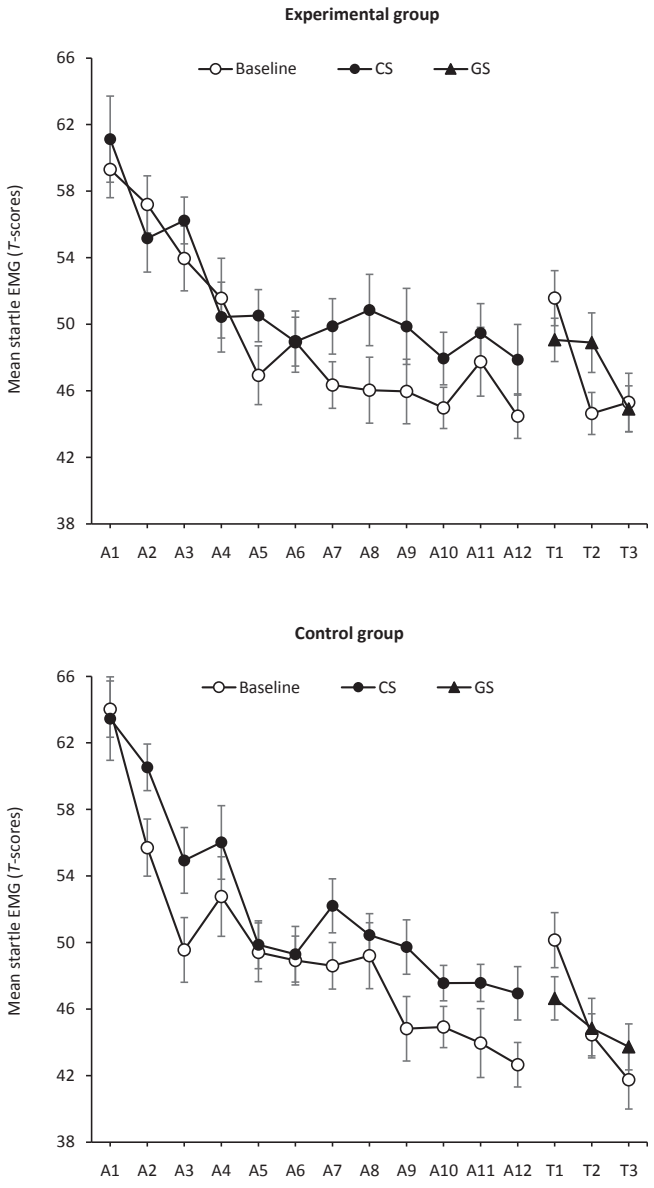

Fig. 7. Mean standardized startle blink responses during the acquisition (A1-12) and test phase (T1-3). Error bars reflect standard errors of the mean.

7. General discussion

We set out to provide objective tests of the effect of EM during memory recall. Experiment I showed that the dual task slows down reaction time in a stimulus discrimination task that requires examination of the image memory. Experiment II demonstrated that EM during recall of a danger cue results in increased generalization of danger expectancies and elevated SCR to stimuli that perceptually resemble the danger cue. These findings are in line with the hypotheses, and provide an extended replication of Van den Hout et al. (2013).

The current study provides additional evidence that the EM manipulation affects memory performance and presents the first demonstration of this effect for hedonically negative stimulus material. These results have several implications. First, our reaction time and skin conductance reactivity data allow for cross-validation of the existing self-report data and the few studies that used physiological measures (e.g., Engelhard et al., 2010; Kearns & Engelhard, 2015). Second, the findings provide additional insight in *how* the EM component adds to EMDR's effectiveness. One account holds that EM during recall causes a permanent loss of stimulus attributes that are represented in memory via updating of the memory trace, i.e. via reconsolidation, which leaves these attributes unavailable for future recall (e.g., Maxfield et al., 2008; Van den Hout & Engelhard, 2012). Such effects are difficult to prove using subjective measures. The present experiments provide objective indexes of memory performance. Yet the interpretation of the results in light of reconsolidation warrants caution. Specifically, it is unclear whether the apparent reduction in memory performance resulted from reduced *accessibility* or *availability* of stimulus attributes (Tulving & Pearlstone, 1966). When stimulus attributes are not accessible, this means that they still exist in long-term memory, but cannot be retrieved. By definition, memory inaccessibility is temporary or context-dependent. In contrast, when stimulus attributes are presumably no longer available, they permanently lose the ability to influence either conscious or unconscious processing (Harris, Sutton, & Barnier, 2010). Important in this regard is that animal research has demonstrated that memory (re)consolidation takes several hours (Nader & Hardt, 2009; Nader, Schafe, & LeDoux, 2000), and behavioral effects may only be expected following a substantial interval (Agren, 2014; see for example; Hupbach, Gomez, Hardt, & Nadel, 2007; Walker, Brakefield, Hobson, & Stickgold, 2003). An explanation of the effects of EM on memory in terms of reconsolidation thus requires the observation of changes in memory availability (such as reduced stimulus discrimination accuracy) following a substantial time delay. To further elucidate whether accessibility or availability of stimulus attributes mediates the effect of EM on memory performance, future research may employ delayed memory tests.

A third implication relates to a potential risk associated with the current EMDR treatment protocol. The starting point in an EMDR session is that the client visualizes the 'hotspot' (i.e., the worst moment of the trauma memory that causes a peak level of emotional distress; e.g., Holmes, Grey, & Young, 2005) while making EMs. As a typical next step, the therapist asks "what comes up" and directs the client to focus on this new (feature of the) memory, and engage in another set of EM (De Jongh & ten Broeke, 2009). For example, in case of a violent attack it is likely that, next to the hotspot (e.g., the moment the knife was put to my neck), contextual stimuli are recalled (e.g., the physical environment or the appearance of the perpetrator). Based on the current data, one may argue that whereas the blurring of emotional hotspots seems conducive, the loss of related (useful) information may actually become a risky endeavor. Specifically, it can be predicted that the blurring of danger cue representations may broaden the range of stimuli capable of evoking a fear response, thereby worsening the psychopathology. Obviously, such an outcome is undesirable and implies that therapists should be alert in determining what memory features to include in the dual task.

Three issues related to this tentative recommendation deserve further attention. First and foremost, the current findings and implications warrant replication. Second, we are not aware of case reports in which EMDR treatment results in stronger fear generalization. Instead, this hypothesis is theory-driven. Note that such an approach may provide valuable insights. For example, over the years, many therapists have replaced EM by other forms of bilateral stimulation, such as listening to alternating beeps (Maxfield, 2008). However, following the advances in theory, it was postulated by van den Hout, Van den Hout and colleagues (2011) that beeps may be less effective, because they hardly use working memory resources. This hypothesis was tested and confirmed, first in a student sample (Van den Hout et al., 2011), and subsequently in a sample of PTSD patients (Van den Hout et al., 2012). Dutch EMDR

trainers now encourage therapists to generally use EM rather than beeps (Beer et al., 2011). A third issue concerns the theoretical distinction between danger cues and hotspots, which may not be obvious in clinical cases. For example, the act of looking the perpetrator in the eyes, rather than the assault itself, may well become the image that causes peak levels of distress. In this example, the perpetrators' appearance may be conceptualized a danger cue. However, in reality, the image does not function as a cue that activates the hotspot memory image, but rather represents the hotspot itself. Accordingly, it may be expected that blurring of the perpetrators' memory representation does not result in increased fear generalization.

A final implication of the current findings is that they indirectly suggest that increasing the vividness and detail of a CS representation will reduce fear generalization. In support of this idea, animal studies have documented that exposure to the training context (CS) prior to context-shock conditioning, which arguably strengthens the context's representation, reduces generalization of conditional freezing in a (novel) test context (Biedenkapp & Rudy, 2007). Also, 'pre-test cuing treatment', i.e. re-exposure to the training context just before entering a (novel) test context, which likely reactivates or reinstates the training context memory, has been shown to reduce generalization of conditional freezing (Wiltgen & Silva, 2007; Zhou & Riccio, 1994). Future research may elucidate whether such pre-test cuing treatment in humans, e.g. via guided imagination, has similar effects.

An unexpected finding was that EM affected generalization of shock expectancy and SCRs, but not of startle responding. There may be several explanations. First, some scholars have identified shock expectancy and SCR primarily as measures of contingency learning, and startle responding as a more specific/affective measure of fear learning (e.g., Soeter & Kindt, 2010). Possibly, then, the current results reflect an increase in expectancy generalization rather than fear generalization. However, recent data show that SCR tracks learning about the estimated UCS-intensity – over and above the estimated CS-UCS contingency – which arguably reflects threat learning (Haesen & Vervliet, 2015). Alternatively, the effect may be short-lived. Whereas SCRs typically rise 1–3 s following stimulus onset (Dawson et al., 2007), and shock expectancy was rated 4066 ms on average after image onset, startle probes were delivered only 4–5 s after stimulus onset. Shock-expectancy was still increased during the second and third test trial, however, which is at odds with this hypothesis. Finally, statistical power may have been too low to detect between-group differences in startle responding. Note that only 36 participants provided clean EMG data, compared with 48 participants providing clean SCR data. Exploratory analyses revealed that the relative increase in generalized responding during the second test trial was associated with a medium effect size for both measures (SCR: $d = 0.63$; startle EMG: $d = 0.51$).

A limitation of this study is that we did not include a measure of behavioral avoidance. Alongside subjective experience and physiology, behavior is core part of the emotional (fear) response (e.g., Mauss & Robinson, 2009). Because avoidance behavior is one of the diagnostic criteria for many anxiety-related disorders (including PTSD), and captures a unique component of the fear response, its measurement is worthwhile (Beckers, Krypotos, Boddez, Effting, & Kindt, 2013). Future research that aims to replicate and extend the current findings may therefore include, for example, an approach-avoidance task (AAT; e.g., Krypotos, Effting, Arnoudova, Kindt, & Beckers, 2014) or the possibility to avoid UCS-occurrence (e.g., Lommen, Engelhard, & Van den Hout, 2010). Another limitation to both experiments I and II relates to the lack of longer-term follow-up tests. It is unclear whether the current results reflect that the EM manipulation renders memory details (temporarily) inaccessible or (permanently) unavailable. A recommendation for future studies is therefore to present the original stimulus amidst similar stimuli after a substantial retention interval, and test for stimulus discrimination.

In conclusion, the current experiments provide a first conceptual replication of Van den Hout et al. (2013) showing that EM during recall attenuates subsequent memory performance. In addition, we have presented a first demonstration of this effect using a hedonically negative stimulus, which benefits its generalization to how EM affects traumatic memories in EMDR. Ultimately, the current findings may contribute to the much needed improvement of EMDR.

Authorship

All authors contributed to the study design. Testing, data collection, data analysis, and drafting of the paper were performed by A.L. All authors provided critical revisions and approved the final version of the paper for submission.

Acknowledgements

This study was supported by a grant from the Netherlands Ministry of Defense (23.93.032.02.52) awarded to Iris M. Engelhard, and by a short stay PhD fellowship from Utrecht University (SSF-Leer 2014) to Arne Leer to conduct this study at the University of Leuven. We thank Titia de Bie and Lynn Hendrikx for their assistance with data collection and Jeroen Clarysse and Riet Fonteyne for their technical support.

References

Agren, T. (2014). Human reconsolidation: A reactivation and update. *Brain Research Bulletin, 105*, 70–82. http://dx.doi.org/10.1016/j.brainresbull.2013.12.010.

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

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*, 209–223. http://dx.doi.org/10.1111/j.2044-8260.1997.tb01408.x.

Bahrick, H. P., Clark, S., & Bahrick, P. (1967). Generalization gradients as indicants of learning and retention of a recognition task. *Journal of Experimental Psychology, 75*, 464–471. http://dx.doi.org/10.1037/h0025131.

Barrowcliff, A. L., Gray, N. S., Freeman, T. C. A., & MacCulloch, M. J. (2004). Eye-movements reduce the vividness, emotional valence and electrodermal arousal associated with negative autobiographical memories. *The Journal of Forensic Psychiatry & Psychology, 15*, 325–345. http://dx.doi.org/10.1080/14789940410001673042.

Beckers, T., Krypotos, A.-M., Boddez, Y., Effting, M., & Kindt, M. (2013). What's wrong with fear conditioning? *Biological Psychology, 92*, 90–96. http://dx.doi.org/10.1016/j.biopsycho.2011.12.015.

Beer, R., ten Broeke, E., Hornsveld, H., de Jongh, A., Meijer, S., de Roos, C., et al. (2011). *EMDR: Oogbewegingen of een andere duale taak?*. Retrieved from http://www.emdr.nl/wp-content/uploads/2016/01/EMDR-Richtlijnen-door-trainers.pdf.

Biedenkapp, J. C., & Rudy, J. W. (2007). Context preexposure prevents forgetting of a contextual fear memory: Implication for regional changes in brain activation patterns associated with recent and remote memory tests. *Learning & Memory, 14*, 200–203. http://dx.doi.org/10.1101/lm.499407.

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

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology, 42*, 1–15. http://dx.doi.org/10.1111/j.1469-8986.2005.00271.x.

Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry, 162*, 214–227. http://dx.doi.org/10.1176/appi.ajp.162.2.214.

Christman, S. D., Garvey, K. J., Propper, R. E., & Phaneuf, K. A. (2003). Bilateral eye movements enhance the retrieval of episodic memories. *Neurophysiology, 17*, 221–229. http://dx.doi.org/10.1037/0894-4105.17.2.221.

Craske, M. G., Hermans, D., & Vansteenwegen, D. (2006). *Fear and learning: From basic processes to clinical implications*. Washington DC: APA books.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159–181). Cambridge, UK: Cambridge University Press.

De Clercq, A., Verschuere, B., De Vlieger, P., & Crombez, G. (2006). Psychophysiological analysis (PSPHA): A modular script-based program for analyzing psychophysiological data. *Behavior Research Methods, 38*, 504–510. http://dx.doi.org/10.3758/BF03192805.

De Jongh, A., & ten Broeke, E. (2009). *Handbook EMDR: Een geprotocolleerde behandelmethode voor de gevolgen van psychotrauma* (4th ed.). Amsterdam: Pearson Assessment and Information.

Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear Generalization in humans: Systematic review and implications for anxiety disorder research. *Behavior Therapy*. http://dx.doi.org/10.1016/j.beth.2014.10.001.

Engelhard, I. M. (2012). Making science work in mental health care. *Psychotraumatology, 3*, 18740. http://dx.doi.org/10.3402/ejpt.v3i0.18740.

Engelhard, I. M., van Uijen, S. L., & van den Hout, M. A. (2010). The impact of taxing working memory on negative and positive memories. *European Journal of Psychotraumatology, 1*, 5623. http://dx.doi.org/10.3402/ejpt.v1i0.5623.

Engelhard, I. M., van den Hout, M. A., Janssen, W., & van der Beek, J. (2010). Eye movements reduce the vividness and emotionality of "flashforwards". *Behaviour Research and Therapy, 48*, 422–447. http://dx.doi.org/10.1016/j.brat.2010.01.003.

Fonteyne, R., Vervliet, B., Hermans, D., Baeyens, F., & Vansteenwegen, D. (2009). Reducing chronic anxiety by making the threatening event predictable: An experimental approach. *Behaviour Research and Therapy, 47*, 830–839. http://dx.doi.org/10.1016/j.brat.2009.06.011.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*, 567–589. http://dx.doi.org/10.1111/j.1469-8986.1986.tb00676.x.

Grillon, C., Pine, D. S., Baas, J. M. P., Lawley, M., Ellis, V., & Charney, D. S. (2006). Cortisol and DHEA-S are associated with startle potentiation during aversive conditioning in humans. *Psychopharmacology, 186*, 434–441. http://dx.doi.org/10.1007/s00213-005-0124-2.

Gunter, R. W., & Bodner, G. E. (2008). How eye movements affect unpleasant memories: Support for a working-memory account. *Behaviour Research and Therapy, 46*, 913–931. http://dx.doi.org/10.1016/j.brat.2008.04.006.

Haesen, K., & Vervliet, B. (2015). Beyond extinction: Habituation eliminates conditioned skin conductance across contexts. *International Journal of Psychophysiology*. http://dx.doi.org/10.1016/j.ijpsycho.2014.11.010.

Harris, C. B., Sutton, J., & Barnier, A. (2010). Autobiographical forgetting, social forgetting and situated forgetting. In S. Della Sala (Ed.), *Forgetting* (pp. 253–284). Hove: Psychology Press.

Herbert, J. D., Lilienfeld, S. O., Lohr, J. M., Montgomery, R. W., O'Donohue, W. T., 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*, 945–971. http://dx.doi.org/10.1016/S0272-7358(99)00017-3.

Holmes, E. A., Grey, N., & Young, K. A. D. (2005). Intrusive images and "hotspots" of trauma memories in posttraumatic stress disorder: An exploratory investigation of emotions and cognitive themes. *Journal of Behavior Therapy and Experimental Psychiatry, 36*, 3–17. http://dx.doi.org/10.1016/j.jbtep.2004.11.002.

Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning & Memory, 14*, 47–53. http://dx.doi.org/10.1101/lm.365707.

Kavanagh, D. J., Freese, S., Andrade, J., & May, J. (2001). Effects of visuospatial tasks on desensitization to emotive memories. *British Journal of Clinical Psychology, 40*, 267–280. http://dx.doi.org/10.1348/014466501163689.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27. http://dx.doi.org/10.1146/annurev.clinpsy.3.022806.091432.

Kearns, M., & Engelhard, I. M. (2015). Psychophysiological responsivity to script-driven imagery: An exploratory study of the effects of eye movements on public speaking flashforwards. *Frontiers in Psychiatry, 6*, 115. http://dx.doi.org/10.3389/fpsyt.2015.00115.

Krypotos, A.-M., Effting, M., Arnoudova, I., Kindt, M., & Beckers, T. (2014). Avoided by association: Acquisition, extinction, and renewal of avoidance tendencies toward conditioned fear stimuli. *Clinical Psychological Science, 2*, 336–343. http://dx.doi.org/10.1177/2167702613503139.

Kuiken, D., Bears, M., Miall, D., & Smith, L. (2002). Eye movement desensitization reprocessing facilitates attentional orienting. *Imagination, Cognition and Personality, 21*, 3–20. http://dx.doi.org/10.2190/L8JX-PGLC-B72R-KD7X.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition & Emotion, 24*, 1377–1388. http://dx.doi.org/10.1080/02699930903485076.

LeDoux, J. E. (2014). Coming to terms with fear. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 2871–2878. http://dx.doi.org/10.1073/pnas.1400335111.

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*, 231–239. http://dx.doi.org/10.1016/j.jbtep.2012.11.001.

Leer, A., Engelhard, I. M., Dibbets, P., & Van den Hout, M. A. (2013). Dual-tasking attenuates the return of fear after extinction. *Journal of Experimental Psychopathology, 4*, 325–340. http://dx.doi.org/10.5127/jep.029412.

Leer, A., Engelhard, I. M., & van den Hout, M. A. (2014). How eye movements in EMDR work: Changes in memory vividness and emotionality. *Journal of Behavior Therapy and Experimental Psychiatry, 45*, 396–401. http://dx.doi.org/10.1016/j.jbtep.2014.04.004.

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

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., et al. (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy, 43*, 1391–1424. http://dx.doi.org/10.1016/j.brat.2004.10.007.

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*, 285–302. http://dx.doi.org/10.1016/0005-7916(95)00041-0.

Lommen, M. J. J., Engelhard, I. M., & Van den Hout, M. A. (2010). Neuroticism and threat avoidance: Better safe than sorry? *Personality and Individual Differences, 49*, 1001–1006. http://dx.doi.org/10.1016/j.paid.2010.08.012.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion, 23*, 209–237. http://dx.doi.org/10.1080/02699930802204677.

Maxfield, L. (2008). Considering mechanisms of action in EMDR. *Journal of EMDR Practice and Research, 2*, 234–238. http://dx.doi.org/10.1891/1933-3196.2.4.234.

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

McNally, R. J. (1999). EMDR and mesmerism: A comparative historical analysis. *Journal of Anxiety Disorders, 13*, 225–236. http://dx.doi.org/10.1016/S0887-6185(98)00049-8.

Nader, K., & Hardt, O. (2009). A single standard for memory: The case for reconsolidation. *Nature Reviews Neuroscience, 10*, 224–234. http://dx.doi.org/10.1038/nrn2590.

Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Reply — reconsolidation: The labile nature of consolidation theory. *Nature Reviews Neuroscience, 1*, 216–219. http://dx.doi.org/10.1038/35044580.

National Institute for Health and Care Excellence. (2005). Post-Traumatic Stress Disorder (PTSD): The management of PTSD in adults and children in primary and secondary care. *NICE clinical guideline, 26*. Retrieved from http://guidance.nice.org.uk/CG26.

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology, 46*, 984–995. http://dx.doi.org/10.1111/j.1469-8986.2009.00852.x.

Riccio, D. C., Rabinowitz, V. C., & Axelrod, S. (1994). Memory: When less is more. *American Psychologist, 49*, 917–926. http://dx.doi.org/10.1037/0003-066X.49.11.917.

Scapinello, K. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21*, 329–330. http://dx.doi.org/10.3758/BF03335807.

Schubert, S. J., Lee, C. W., & Drummond, P. D. (2011). The efficacy and psychophysiological correlates of dual-attention tasks in eye movement desensitization and reprocessing (EMDR). *Journal of Anxiety Disorders, 25*, 1–11. http://dx.doi.org/10.1016/j.janxdis.2010.06.024.

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. http://dx.doi.org/10.1017/S0033291706007963.

Shapiro, F. (2001). *Eye movement desensitization and reprocessing: Basic principles, protocols and procedures* (2nd ed.). New York: Guilford Press.

Soeter, M., & Kindt, M. (2010). Dissociating response systems: Erasing fear from memory. *Neurobiology of Learning and Memory, 94*, 30–41. http://dx.doi.org/10.1016/j.nlm.2010.03.004.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologist Press.

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A free software package for implementing psychological and psychophysiological experiments. *Experimental Psychology, 57*, 36–45. http://dx.doi.org/10.1027/1618-3169/a000005.

Struyf, D., Zaman, J., Vervliet, B., & Van Diest, I. (2015). Perceptual discrimination in fear generalization: Mechanistic and clinical implications. *Neuroscience & Biobehavioral Reviews, 59*, 210. http://dx.doi.org/10.1016/j.neubiorev.2015.11.004, 207.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381–391. http://dx.doi.org/10.1016/S0022-5371(66)80048-8.

Van Schie, K., Engelhard, I. M., & van den Hout, M. A. (2015). Taxing working memory during retrieval of emotional memories does not reduce memory accessibility when cued with reminders. *Frontiers in Psychiatry, 6*, 16. http://dx.doi.org/10.3389/fpsyt.2015.00016.

Van den Hout, M. A., Bartelski, N., & Engelhard, I. M. (2013). On EMDR: Eye movements during retrieval reduce subjective vividness and objective memory accessibility during future recall. *Cognition & Emotion, 27*, 177–183. http://dx.doi.org/10.1080/02699931.2012.691087.

Van den Hout, M. A., Eidhof, M. B., Verboom, J., Littel, M., & Engelhard, I. M. (2014). Blurring of emotional and non-emotional memories by taxing working memory during recall. *Cognition and Emotion, 28*, 717–727. http://dx.doi.org/10.1080/

02699931.2013.848785.
Van den Hout, M. A., & Engelhard, I. M. (2012). How does EMDR work? *Journal of Experimental Psychopathology, 3*, 724–738. http://dx.doi.org/10.5127/jep.028212.
Van den Hout, M. A., Engelhard, I. M., Rijkeboer, M. M., Koekebakker, J., Hornsveld, H., Leer, … Akse, N. (2011). EMDR: Eye movements superior to beeps in taxing working memory and reducing vividness of recollections. *Behaviour Research and Therapy, 49*, 92–98. http://dx.doi.org/10.1016/j.brat.2010.11.003.
Van den Hout, M. A., Engelhard, I. M., Smeets, M. A. M., Hornsveld, H., Hoogeveen, E., de Heer, E., … Rijkeboer, M. (2010). Counting during recall: Taxing of working memory and reduced vividness and emotionality of negative memories. *Applied Cognitive Psychology, 24*, 1–9. http://dx.doi.org/10.1002/acp.1677.
Van den Hout, M., Muris, P., Salemink, E., & Kindt, M. (2001). Autobiographical memories become less vivid and emotional after eye movements. *British Journal of Clinical Psychology, 40*, 121–130. http://dx.doi.org/10.1348/014466501163571.
Van den Hout, M. A., Rijkeboer, M. M., Engelhard, I. M., Klugkist, I., Hornsveld, H., Toffolo, M. B. J., et al. (2012). Tones inferior to eye movements in the EMDR treatment of PTSD. *Behaviour Research and Therapy, 50*, 275–279. http://dx.doi.org/10.1016/j.brat.2012.02.001.
Vervliet, B., & Geens, M. (2015). Fear generalization in humans: Impact of feature learning on conditioning and extinction. *Neurobiology of Learning and Memory, 113*, 143–148. http://dx.doi.org/10.1016/j.nlm.2013.10.002.
Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010). Fear generalization in humans: Impact of verbal instructions. *Behaviour Research and Therapy, 48*, 38–43. http://dx.doi.org/10.1016/j.brat.2009.09.005.
Walker, M. P., Brakefield, T., Hobson, A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature, 425*, 616–620. http://dx.doi.org/10.1038/nature01930.
Wiltgen, B. J., & Silva, A. J. (2007). Memory for context becomes less specific with time. *Learning & Memory, 14*, 313–317. http://dx.doi.org/10.1101/lm.430907.
Zhou, Y. L., & Riccio, D. C. (1994). Pretest cuing can alleviate the forgetting of contextual stimulus attributes. *Learning & Motivation, 25*, 233–244. http://dx.doi.org/10.1006/lmot.1994.1013.