

## Predictors of involuntary and voluntary emotional episodic memories of virtual reality scenarios in Veterans with and without PTSD

Loretta S. Malta, Cezar Giosan, Lauren E. Szkodny, Margaret M. Altemus, Albert A. Rizzo, David A. Silbersweig & JoAnn Difede

To cite this article: Loretta S. Malta, Cezar Giosan, Lauren E. Szkodny, Margaret M. Altemus, Albert A. Rizzo, David A. Silbersweig & JoAnn Difede (2020): Predictors of involuntary and voluntary emotional episodic memories of virtual reality scenarios in Veterans with and without PTSD, *Memory*, DOI: [10.1080/09658211.2020.1770289](https://doi.org/10.1080/09658211.2020.1770289)

To link to this article: <https://doi.org/10.1080/09658211.2020.1770289>



Published online: 28 May 2020.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



## Predictors of involuntary and voluntary emotional episodic memories of virtual reality scenarios in Veterans with and without PTSD

Loretta S. Malta<sup>a</sup>, Cezar Giosan <sup>a,b</sup>, Lauren E. Szkodny<sup>a</sup>, Margaret M. Altemus<sup>a</sup>, Albert A. Rizzo <sup>a,c</sup>, David A. Silbersweig<sup>a</sup> and JoAnn Difede<sup>a</sup>

<sup>a</sup>Weill Medical College of Cornell University, 1300 York Ave, New York, NY 10065; <sup>b</sup>Department of Psychology, University of Bucharest, Panduri 90, Bucharest, Romania, 032075; <sup>c</sup>Institute for Creative Technologies, University of Southern California, 12015 Waterfront Drive, Playa Vista, CA 90094–2536

### ABSTRACT

This study investigated predictors of involuntary and voluntary memories of stressful virtual reality scenarios. Thirty-two veterans of the two Persian Gulf Wars completed verbal memory tests and diagnostic assessments. They were randomly assigned to a Recounting (16) or a Suppression (16) condition. After immersion in the VR scenarios, the Recounting group described the scenarios and the Suppression group suppressed thoughts of the scenarios. One week later, participants completed surprise voluntary memory tests and another thought suppression task. The best predictors of voluntary memory were verbal memory ability, dissociation, and to a lesser extent, physiological arousal before and after scenarios. Dissociation and physiological stress responses selectively affected memory for neutral elements. Higher distress during scenarios impaired voluntary memory but increased the frequency of involuntary memories. Physiological stress responses promoted more frequent involuntary memories immediately after the scenarios. More frequent initial involuntary memories, tonic physiological arousal, and stronger emotional responses to dangerous events predicted difficulty inhibiting involuntary memories at follow-up. The effects of thought suppression were transient and weaker than those of other variables. The findings suggest that posttraumatic amnesia and involuntary memories of adverse events are more related to memory ability and emotional and physiological stress responses than to post-exposure suppression.

### ARTICLE HISTORY

Received 9 August 2019  
Accepted 11 May 2020

### KEYWORDS

PTSD; memory; virtual reality; stress

### Introduction

This study investigated mechanisms that influence the development of memories of adverse events. Trauma survivors frequently think about traumatic events in their early aftermath (McMillen et al., 2000), but how this develops into persistent re-experiencing symptoms is not well understood. Moreover, repeated involuntary trauma memories and posttraumatic amnesia are both symptoms of posttraumatic stress disorder (PTSD) (American Psychiatric Association, 2013), and there is continued debate over whether PTSD is a disorder of abnormal remembering (Brewin, 2014) or abnormal forgetting (Berntsen & Rubin, 2014).

One mechanism that could promote both symptoms is memory suppression (Abramowitz et al., 2001; Anderson & Levy, 2009; Sullivan et al., 2019; Wenzlaff & Wegner, 2000). Suppression, compared to memorisation or passive viewing, impairs cued recall of paired associates (Anderson & Levy, 2009) and memories of recent events (Hulbert et al., 2016). Studies that have used films as stimuli found that suppression did not impair memory for film content

(Nixon et al., 2009; Rassin et al., 1997; Wegner et al., 1996), but one study (Wegner et al., 1996) found that it did impair recognition memory for the chronology of events in the film. The effect failed to replicate (Nixon et al., 2009; Rassin et al., 1997; Wegner et al., 1996), although Wegner et al.'s (1996) second experiment found that suppression predicted poorer free recall of the sequence of events in the film.

Trauma survivors are not necessarily unable to recall the order of events, but may report memory gaps between events (Brewin, 2014; Brewin et al., 2010; Sachschal et al., 2019). This could be due to dissociating during trauma exposure (Bergouignan et al., 2014; Brewin, 2011), but difficulty evoking dissociation in laboratory studies has led to inconsistent findings (Holmes & Bourne, 2008).

Paradoxically, suppression increases the frequency of involuntary thoughts during subsequent periods of unrestricted thinking (Abramowitz et al., 2001; Wenzlaff & Wegner, 2000). The modest magnitude of this “rebound” effect has led researchers to propose that other processes must be involved (Abramowitz et al., 2001). Candidate

mechanisms include differences in suppression ability (Brewin & Smart, 2005; Catarino et al., 2015; Levy & Anderson, 2008), but there is little research testing whether this affects the development of involuntary episodic memories. Another open domain of inquiry is the effect of pre-exposure psychopathology. PTSD diagnosis has not consistently affected suppression ability or rebound effects (Beck et al., 2006; Catarino et al., 2015; Shipherd & Beck, 1999) and there is scant research on the effects of specific PTSD symptoms. Hyperarousal symptoms (startle, hypervigilance, poor sleep and concentration, irritability) may be predisposing, rather than acquired, symptoms (Admon et al., 2013; Gehrman et al., 2013; van Liempt et al., 2013). Hyperarousal symptoms could index dysregulation in brain regions involved in emotional learning and memory (Admon et al., 2013) and this could hypothetically lead to persistent, intrusive trauma memories (LeDoux, 2000).

One theory of PTSD posits that hyperarousal enhances perceptual trauma memories but impairs episodic memory (Brewin, 2011, 2014; Brewin et al., 2010). The vividness of the perceptual memories and their weak integration with contextual features renders the memories more easily activated by cues and thus more intrusive. Deficient episodic memories manifest as re-experiencing events outside of their temporal-spatial context (flashbacks) and amnesia. Thus, processes in two distinct memories systems purportedly lead to both involuntary memories and posttraumatic amnesia. Accordingly, this “dual process” theory predicts either no correlation or a negative correlation between the accuracy of voluntary recall and the frequency of involuntary memories (Brewin, 2014). A contrasting theory posits that PTSD reflects abnormal forgetting (Berntsen & Rubin, 2014) and that persistent involuntary memories develop from trauma memories that are too easily recalled. This “inability to forget” theory predicts a positive correlation between the ease of voluntary recall and the frequency of involuntary memories. This hypothesis has been supported when memory frequency is assessed with a rating scale (Berntsen & Rubin, 2014). However when subjects log the actual occurrence of involuntary memories, correlations with the accuracy of cued recall are either non-significant or negative (Brewin, 2014). The diary method might yield more accurate frequency estimates, but monitoring could increase the frequency of involuntary memories (Barzykowski & Niedźwieńska, 2016; Barzykowski & Staugaard, 2018; Barzykowski, Niedźwieńska, et al., 2019; Wenzlaff & Wegner, 2000) and might affect voluntary memory as well. Research testing these theories has been conducted primarily with non-clinical samples (Brewin, 2014). This is problematic because PTSD is associated with differential performance on memory tests (Guez et al., 2011; Marx et al., 2009) and reduced volume of brain regions involved in memory that could reflect predisposing risk factors (Admon et al., 2013; Karl et al., 2006).

The aim of present study was to advance our understanding of emotional, physiological, and mental responses of veterans who may have been exposed to combat trauma, so as to better understand how these responses might promote the development of posttraumatic stress symptoms. By better understanding how symptoms develop, we may be able to create more effective prevention programmes and improve current treatments.

To this end, the authors created a novel virtual reality (VR) research paradigm to prospectively investigate variables that might influence the development of involuntary and voluntary episodic memories. While virtual reality simulations have been used in exposure therapy for anxiety disorders in the past (Pause et al., 2013), they have been underutilised in experimental psychopathology. Virtual reality environments are ideal for studying episodic emotional memories because events can be standardised, can unfold in real time and be experienced from a first-person perspective, thus evoking the egocentric, narrative and temporal aspects of stressful events (Bergouignan et al., 2014; Pause et al., 2013). The use of standardised scenarios avoids limitations associated with studying actual trauma memories, the reliability of which is influenced by the severity of PTSD at the time of the assessment (Giosan et al., 2009). Memory reconsolidation is subject to interference from various factors (Schwabe et al., 2014; Vallejo et al., 2019), therefore, by standardising stimuli that can elicit episodic memories, we can better pinpoint the mechanisms behind these phenomena. This, in turn, may lead to better prevention and treatment programmes.

To this end, in the present study participants were immersed in VR scenarios as their stress responses were assessed, followed by either recollecting the scenarios or suppressing memories of them. We hypothesised that suppression would increase the frequency of involuntary memories and impair the ability to recall the sequences of events. We tested the competing hypotheses that the frequency of involuntary memories would correlate with either impaired or enhanced voluntary episodic memory. We also hypothesised that PTSD symptom severity, weaker thought inhibition ability, and stronger emotional and physiological reactions to the VR scenarios would promote more frequent involuntary memories; and that verbal memory ability and the intensity of stressor responses during the scenarios would predict the accuracy of voluntary memories.

## Method

### Participants

The study was approved by the Institutional Review Board of Weill Medical College of Cornell University.<sup>1</sup> All participants signed a consent form, which informed of the potential benefits and possible risks of the study (e.g.,

experiencing distress, possible increase of PTSD symptoms, potential motion sickness, potentially distressing questions in the psychological evaluation).

Participants were Veterans of the two Persian Gulf Wars (N=5 served in the Gulf War 1, 8/2/1990–2/28/1991, months since returning from warzone: min = 162 – max = 192; N=25 were Iraq war vets, 3/20/03–12/18/11, months since returning from warzone: min = 2 – max = 96; N=2 were vets who served in both wars, months since returning from warzone: min = 5 – max = 10 months; Data were collected between 2006–2008. Recruitment started on 9/15/2006. Last person enrolled was on 8/5/2008).

Exclusions: psychosis, cognitive impairment, substance dependence, Attention Deficit Hyperactivity Disorder (ADHD), motion sickness, conditions that prevented standing. Fifty Veterans completed the diagnostic evaluation to determine eligibility, 36 were enrolled, and four dropped out after the diagnostic evaluation. The final sample consisted of trauma-exposed Veterans with varying levels of PTSD symptoms. Eleven participants (34%) were diagnosed with PTSD of moderate severity, Mean CAPS total = 59.82,  $SD = 10.23$ ; and they endorsed mild symptoms of dissociation, Mean CAPS Dissociation score = 3.09,  $SD = 3.39$ . The remainder of the sample presented with mild symptoms of PTSD, Mean CAPS total = 17.05,  $SD = 13.15$ ; and absent/ minimal symptoms of dissociation, Mean CAPS Dissociation score = 0.81,  $SD = 1.99$ .

The participants were randomly assigned to a Recounting (16) or a Suppression (16) condition. There were 27 males and five females aged 21–46 years old, with a mean ( $SD$ ) age of 30.28 (7.32). Racial/ethnic minorities comprised 56.25% of the sample. The majority (50%) had completed one-two years of college, 12.50% were high school graduates, and 37.50% were college graduates. Participants provided informed consent and were told that

study risks included simulator sickness (Regan, 1995) and a transient increase in PTSD symptoms. They were compensated \$50.00/visit.

### Overview of procedures

The study was conducted in three visits scheduled one week apart: a diagnostic evaluation, the VR challenge, and follow up testing. The diagnostic evaluation consisted of clinical interviews, symptom questionnaires, verbal memory tests, and sampling of salivary cortisol. Participants with mental health conditions were provided with referrals. Those enrolled in the study were matched according to age, gender, and PTSD severity and randomly assigned to a Suppression or Recounting condition. At the second visit, participants completed a VR challenge task, in which they were shown VR combat scenarios as heart rate, cortisol secretion, and distress level were measured. After the scenarios, the Recounting group completed a free recall test of memory for the scenarios and the Suppression group suppressed thoughts of the scenarios during a thought suppression task. One week later, all participants completed surprise voluntary memory tests, another thought suppression task, and a test of verbal inhibition ability (see Figure 1).

### Instruments

Diagnostic instruments were selected based on their demonstrated reliability and validity in research during the past twenty years.

The Clinician Administered PTSD Scale for DSM-IV symptoms of PTSD (Blake et al., 1998) is an interview that provides a categorical diagnosis and dimensional symptom scores. Symptoms are assigned on a 0–4 scale. Scores for individual symptoms are summed to create subscales for

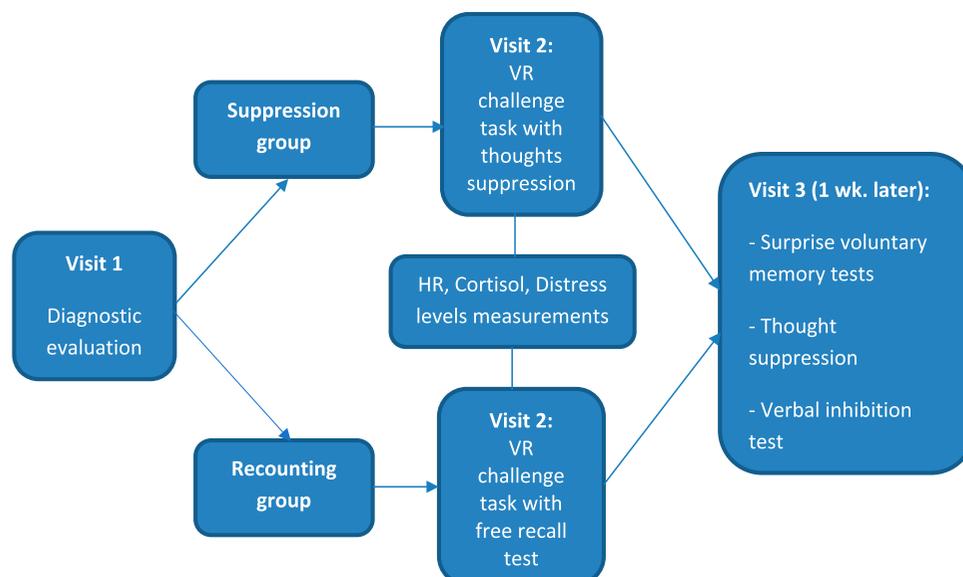


Figure 1. Method.

Re-experiencing, Avoidance/numbing, and Hyperarousal symptom subscales. PTSD diagnosis present during the past month was diagnosed according to recommended CAPS scoring rules (Weathers et al., 1999) and DSM-IV criteria (American Psychiatric Association, 1994). Reliability checks conducted by an independent evaluator on a random selection of 25% of assessments yielded a kappa coefficient of 1.0,  $p < .000$ . The CAPS also assesses symptoms of dissociation (reduced awareness), depersonalisation (reduced body sense), and derealization (altered perception). The CAPS does not have a subscale for these items, but for analyses in the present study we followed the instrument's procedures for creating subscales and summed the three item scores to create a monthly dissociation index.

The PTSD checklist (Weathers et al., 1993) is a self-report questionnaire that assessed DSM-IV symptoms of PTSD (American Psychiatric Association, 1994). The PCL was administered at each visit to assess weekly symptom severity. Items are summed to obtain a total score. To normalise kurtotic score distributions, total scores for weeks two and three were recoded according to percentile rank.

The Structured Clinical Interview for DSM-IV Axis I Disorders (First et al., 1997) provides categorical diagnoses of DSM-IV disorders. The ADHD Interview/Checklist (Barkley & Murphy, 2005) is a checklist administered in an interview format to obtain a diagnosis according to DSM-IV criteria.

The Beck Depression Inventory-II (Beck et al., 1996) is a 21-item scale that provides a dimensional score of depression symptoms.

Verbal memory was assessed with the NYU Memory Test (Kluger et al., 1999), a test of immediate and delayed free recall of paragraphs about events and memory for word pairs. Scores on this test significantly correlate with hippocampal volumes (Golomb et al., 1996).

Verbal inhibition ability was assessed with the Stroop Color-Word Test (Golden & Freshwater, 2002), a well-known test that requires participants to name the colour of the ink in which words are printed while suppressing competing colour word names. Verbal inhibition ability was assessed with the paper version of the Stroop Color-Word Test (Golden & Freshwater, 2002; Trenerry et al., 1989). The stimuli are colour names (red, blue, green) printed either in a congruent colour, (the word "red" printed in red ink), or in a different ink, (the word "red" printed in blue ink), which creates interference by requiring the colour word name to be suppressed. Participants are given 45 s to name the colour ink in which each word is printed. The Color-Word Interference T-score is obtained as follows: First, the number of correctly named ink colors are summed to create an uncorrected Color-Word Score. This score is adjusted for age and education level using norms provided in the manual (Golden & Freshwater, 2002) to provide a Predicted Color-Word Score. An Interference Raw Score is calculated by subtracting the Predicted Color-Word Score from the uncorrected Raw Color-Word

Score. The Color-Word Interference T-score is obtained from norms provided in the manual.

### **VR apparatus and scenarios**

Simulations were presented with the Virtually Better System (Decatur, GA: Virtually Better, 2005), which included a Dell computer and monitors, a head-mounted display with headphones and a head movement tracker, and a platform on which the user stood that vibrated with explosions and other auditory effects. Simulations immersed users in a three-dimensional, multi-sensory environment in which imagery changed in a natural way with head and body motion.

The VR scenarios were created from the Iraq World virtual environment (Rizzo et al., 2006) developed to treat war-related PTSD. Scenario 1 was an automated, 2-minute patrol of an Iraq city. Participants were exposed to a rocket-propelled grenade attack, machine gun fire, and improvised explosive devices. Scenario 2 consisted of a 5.5-minute narrative scripted into five, 45–90 s segments. Participants were given navigation instructions before each segment and used a keypad to navigate the environment. During the first two segments they patrolled an Iraq city. During the next two segments an attack with machine guns and bombing occurred. The attack ended during the last segment and the patrol was resumed. Neither scenario contained graphic injuries or deaths.

### **VR challenge, memory tasks, and follow-up visit**

Participants provided a saliva sample and completed the PCL (Weathers et al., 1993). During the 5-min. measurement of baseline heart rate, participants stood and remained as still as possible. (See below for assessment of heart rate and salivary cortisol). Participants then put on the head-mounted display (HMD) and were immersed in the first VR scenario. Upon its completion, participants removed the HMD, provided Subjective Units of Distress ratings (described below) and completed a questionnaire (described below) to assess immersion in Scenario 1. Then they put on the HMD and were immersed in Scenario 2. Participants remained in the HMD while providing SUDS ratings during this scenario. Upon its completion, they removed the HMD and completed the questionnaire to assess immersion in this scenario. Then they completed either a thought suppression task (Suppression Condition) or a free recall memory test (Recounting Condition). Participants in the Suppression Condition suppressed thoughts of the scenarios for 3 min. They were instructed to press a computer key (which recorded thought frequency in a data file) each time an involuntary memory of the scenarios occurred, and then to immediately suppress the thought by thinking about a white table. Participants completed a questionnaire that assessed suppression strategies, thought content, and ratings (0-100 scale) of the effort exerted to suppress thoughts.

Participants in the Recounting Condition completed a free recall test of memories for the scenarios as their responses were recorded, followed by the provision of ratings (0-100 scale) of the amount of effort exerted to recall scenarios. All participants provided another saliva sample and were told that they would return the following week for “testing of mental abilities”. To avoid creating demand characteristics, they were not told about the memory or suppression tests or instructed to suppress thoughts or to avoid talking about the scenarios or asked to log the frequency of involuntary thoughts.

At the follow-up visit, participants completed a PCL (Weathers et al., 1993) and the following tasks (listed in order of completion): free recall and multiple choice tests of memory of the scenarios, thought suppression task, task questionnaires, weekly thought frequency questionnaires, and the Stroop test. They were debriefed and asked about their experiences completing the study. One week later, they were telephoned to assess for any distress related to study participation.

### ***Assessments of immersion and emotional distress during scenarios***

The Immersion/Presence Questionnaire (IPQ) was adapted from instruments used in VR research (Witmer & Singer, 1998; Zimand et al., 2001) and consisted of seven items endorsed on a 7-point scale. Four items that assessed emotional engagement, sensory-perceptual engagement, ability of simulations to evoke feelings similar to real events, and distraction were summed to create an Immersion subscale. The remaining three items that assessed time perception, losing track of events, and emotional numbing were summed to create a Dissociation subscale. Mean Immersion and Dissociation scores were averaged across the two scenarios to create one variable for each subscale. The Dissociation subscale score was recoded (1 = none, 2 = mild, 3 = mild-to-moderate, 4 = moderate to marked) to normalise a kurtotic score distribution.<sup>2</sup>

Distress during the scenarios was assessed with Subjective Units of Distress Scale (SUDS) ratings (0-100 scale). For Scenario 1, pre- and post-scenario and a retrospective estimate of peak SUDS ratings were obtained. Scenario 2 ratings were obtained before and after each of the five segments.

### ***Assessment of involuntary and voluntary memories of the scenarios***

Two dependent variables were created for the frequency of thoughts during the thought suppression tasks: TS1 for the task completed immediately after the VR challenge and TS2 for the task completed one week later. The TS1 variable was the number of involuntary thoughts. The TS2 variable was recoded into a frequency score (1 = none; 2 = 1-2 thoughts, 3 = 3-5; 4 = 6-9, 5 = 10 or more thoughts) to normalise the score distribution.<sup>3</sup> Data from one participant

were missing due to his not having pressed a key to indicate thought frequency during either task.

The frequency of thoughts about the scenarios during the week after the VR challenge and efforts to suppress the thoughts were assessed with a questionnaire that was administered at the follow up visit. Frequency was estimated using a 5-point scale (0 = not at all, 4 = daily/near daily). The weekly thought frequency variable was recoded for analyses as 1 = not at all, 2 = one-two times, 3 = three or more times to normalise the score distribution (See Note 3).

Procedures for scoring the immediate and delayed free recall test of memories for the scenarios were as follows. Two raters, who were unaware of assignment to experimental condition, independently created and scored written transcriptions of the recorded recollections and then compared results to arrive at a final consensus. One point was assigned for each object or event recalled. Additional points were given for recall of descriptive details (e.g., colour). Scores for recall of event sequences were calculated by assigning one point for each pair of correctly recalled consecutive events. Scores were averaged for the two scenarios to create one variable each for immediate and delayed recall of danger-related events, neutral events, danger-related stimuli (objects, people), neutral stimuli, and event sequences. The delayed recall of danger-related events variable was recoded with percentile scores (1 = at/below 25th percentile, 2 = at/below 50<sup>th</sup>, 3 = at/below 75<sup>th</sup>, 4 = above 75<sup>th</sup>) to normalise the distribution.<sup>4</sup> To test overall delayed free recall memory between the Suppression and Recounting Groups, three summary variables were created: Total Delayed Free Recall Memory (sum of all stimuli, events, and event sequences); Total Delayed Free Recall Memory - Neutral (sum of all neutral stimuli and events); and Total Delayed Free Recall Memory - Danger (sum of all danger stimuli and the raw score for danger-related events).

Verbal recognition memory for the scenarios was assessed with a 40-item multiple choice test. The first 14 items were about Scenario 1 and 26 items were about Scenario 2. Half of the items were danger-related, and half were neutral. How strongly the item was associated with danger in the scenarios was rated, on a 0-4 scale, by evaluators who were unaware of the designated item valence. Danger-related items were rated as more strongly associated with danger than neutral items,  $F(1,38) = 32.095$ ,  $p < .001$ ; Mean (SD) ratings = 2.38 (1.19) and 0.66 (0.63) for danger and neutral items, respectively. Two dependent variables were created: percent of correctly recalled neutral items and percent of correctly recalled danger-related items.

### ***Assessment of salivary cortisol amylase***

Visits were scheduled between 12:30-4:30 PM. Participants were instructed to avoid ingesting food or caffeinated beverages, smoking, or exercising one hour before the

appointment. Saliva was collected with salivettes placed beneath the tongue for five minutes (Newton, NC: Sarstedt, Incorporated). Samples were collected at the beginning of each of the first two visits. A third sample was collected at the second visit, after the VR challenge was completed. This sample was collected a minimum of 20 min. after the beginning of the VR challenge (start of Scenario 1) to ensure adequate time for cortisol to manifest in saliva (Kirschbaum et al., 1993). We tried to standardise the duration of time between the beginning of the VR challenge and the post-challenge saliva collection to 30 min. However, due to variability in administration time, although the Mean (SD) duration was 30.41 (4.51) min., the range was 20–43 min. The time of 43 min was due to one outlier whose time was greater than two SD above the mean. With this score removed, the new Mean (SD)=30.41 (4.51) min. and the new range was 20–38 min.<sup>5</sup> Samples were frozen at  $-80^{\circ}$  C until assayed using a commercial ELISA assay optimised for saliva samples (Salimetrics LLC, State College, PA) and then thawed and centrifuged to remove particles. Samples were run in duplicate and the mean value was used for raw data analyses. Intra-assay variability was less than 10% for all samples. Four dependent variables were created: cortisol level at Visit 1, before and after the VR challenge, and pre-to-post challenge change scores. All variables except the change score were recoded according to raw score percentile ranking value to normalise score distributions (See Note 4).

### Assessment of heart rate

Continuous heart rate was assessed with the ambulatory Life Vest System (Ventura, CA: Vivometrics, 1999–2004) which consists of a vest outfitted with three standard electrocardiogram leads that attach to disposable dry electrodes and a portable data collection unit. Data were exported to the CardioEdit programme (University of Illinois at Chicago Brain–Body Center, Director, Stephen W. Porges). Artifacts were removed via visual inspection according to recommended procedures (Inter-beat Interval Training and Reliability Program Manual Version 2, University of Illinois at Chicago Brain–Body Center, 2006–2007). Heart rate variability (HRV) during the baseline period was calculated as the standard deviation of normal-to-normal (NN) intervals, i.e., all intervals between adjacent electrocardiogram QRS complexes resulting from sinus node depolarizations (Malik et al., 1996). HRV data from two participants were excluded due to artifacts greater than 10% of NN values. Mean heart rate was calculated for the following epochs: baseline, Scenario 1, and each of the five segments of Scenario 2. For Scenario 1, scores for change from baseline were calculated by subtracting the mean baseline heart rate from the mean heart rate during Scenario 1. Baseline heart rate and change scores for the two participants with the excessive artifacts during the baseline period were excluded from analyses.

### Data analyses

Experimental Condition was coded as 1 = Suppression and 2 = Recounting for analyses.

Dependent variables for involuntary memories were thoughts during the suppression tasks (TS1 and TS2) and ratings of the frequency of thoughts during the week. Dependent variables for voluntary memory were scores for immediate and delayed recall of stimuli, events, and event sequences; delayed Multiple Choice test scores, and Delayed Recall Summary scores (total, Neutral, Danger). Univariate and Multivariate One-Way Analysis of Variance (ANOVA) and Analysis of Covariance (ANCOVA) tests were used to compare differences between the Experimental Groups on the summary memory variables, and between-group equivalence in sample characteristics and emotional and physiological reactions to the scenarios.

Multiple regression was the primary means of analyses used to test study hypotheses, as we were most interested in the testing the effects of multiple predictors developing a comprehensive model rather than testing the effects of individual variables. Independent variables included Experimental Condition, current mental health symptoms (scores on all CAPS subscales, PCL, and BDI), physiological arousal (cortisol levels, HRV, and heart rate), cognitive abilities (Stroop Interference T-scores, NYU Memory Test scores), and reactions to the VR scenarios (SUDS ratings, Immersion and Dissociation subscale scores). For analyses of predictors of involuntary memories/thoughts of the scenarios, zero-order correlations with voluntary memory scores on the memory tests were also examined. To reduce the number of variables in each analysis, for each dependent variable, we conducted a correlation analysis to identify variables that were significantly correlated with it. We excluded all non-significant variables from further analyses. Significant zero-order variables were entered into a final, stepwise regression analysis for each dependent variable of interest. The researchers did not adjust the Type 1 error significance level to control for multiple analyses because this would have compromised statistical power and increased Type 2 error (Althouse, 2016; Feise, 2002; Rothman, 1990), given the small sample size. Instead, effects sizes are presented and emphasized in the interpretation of results.

## Results

### Sample characteristics

The sample's mean (SD) CAPS total score was 31.75 (23.90). Eleven participants (34%) were diagnosed with PTSD of moderate severity, Mean (SD) CAPS total = 59.82 (10.23); and had mild levels of dissociation, Mean (SD) = 3.09 (3.39). The remainder had mild symptoms, Mean (SD) CAPS total = 17.05 (13.15); and minimal dissociation, Mean (SD) = 0.81 (1.99).

**Table 1.** Reactions to Scenarios: Subjective Units of Distress (SUDS) Ratings and Heart Rate.

	Suppression Condition	Recounting Condition Mean (SD)	Total Sample
<b>SUDS</b>			
Scenario 1, pre	12.50 (13.78)	19.38 (17.78)	15.94 (16.04)
Scenario 1, peak	39.06 (18.28)	49.06 (33.28)	44.06 (26.89)
Scenario 1, post	31.19 (20.87)	41.56 (30.04)	36.38 (25.98)
Scenario 2, pre	21.00 (16.50)	29.38 (27.20)	25.32 (22.69)
Scenario 2, Segments 1 & 2, patrolling for insurgents	27.77 (22.97)	32.94 (31.20)	30.44 (27.20)
Scenario 2, Segments 3 & 4, attack	43.00 (25.59)	46.72 (32.28)	44.92 (28.81)
Scenario 2, post	36.33 (23.79)	46.88 (35.07)	41.77 (30.13)
<b>Heart Rate, Beats Per Minute</b>	84.69 (15.29)	91.10 (14.07)	88.23 (14.72)
Baseline	82.89 (16.05)	88.21 (13.44)	85.73 (14.70)
Scenario 1	85.02 (14.26)	86.05 (12.14)	85.57 (12.95)
Scenario 2, Segment 1	83.78 (14.41)	86.22 (12.50)	85.08 (13.24)
Scenario 2, Segment 2	83.90 (14.56)	86.24 (12.14)	85.15 (13.14)
Scenario 2, Segment 3	82.71 (13.16)	86.51 (12.64)	84.73 (12.80)
Scenario 2, Segment 4			
Scenario 2, Segment 5	84.53 (14.49)	87.12 (12.76)	85.91 (13.42)
Heart Rate Variability during Baseline	0.049 (0.22)	0.049 (0.18)	0.049 (0.20)

The only significant differences between the Recounting and Suppression groups were more racial minorities (i.e., non-Whites) in the Recounting Group,  $\chi^2 = 4.571$ ,  $p = .033$ ; and higher NYU Memory Test paired associates delayed recall scores in the Suppression Group,  $F(2,29) = 5.105$ ,  $p = .032$ . These variables were used as covariates in relevant analyses. All participants denied expecting the follow up tests. They used a variety of suppression strategies in addition to (or other than) thinking about the white table, typically some type of competing mental activity. There were no significant group differences in adherence to instructions and this variable did not predict the frequency of thoughts during the thought suppression tasks. There were no significant group differences in efforts to suppress or recall scenarios.

### Reactions to the scenarios

We conducted a MANOVA for HR, SUDS, Immersion and Dissociation and a MANCOVA (controlling for gender). The results showed that there were no significant group differences in SUDS, Immersion, Dissociation, heart rate, or cortisol levels. Mean (SD) Immersion scores were 18.81 (3.98) for the sample and 18.57 (3.53) and 19.03 (4.46) for the Suppression and Recounting groups, respectively, indicating moderate immersion (mean item score of 4.70). Mean (SD) Dissociation scores were 2.45 (1.09) for the sample and 2.40 (1.05) and 2.50 (1.15) for the Suppression and Recounting groups, respectively, indicating mild dissociation. SUDS ratings are shown in Table 1.

The scenarios evoked moderate distress, with Scenario 1 mean (SD) peak ratings of 44.06 (26.89) and mean (SD) ratings of 44.92 (28.81) for Scenario 2 attack segments. Cortisol level decreased after the VR challenge in 61% of participants, increased in 35%, and did not change in 3%. Mean (SD) change scores were  $-.037$  (.096) ug/dl for the sample,  $-.020$  (.081) ug/dl for the Suppression group,

and  $-.053$  (.108) for the Recounting group. As shown in Table 1, heart rate decreased during Scenario 1 and showed minor fluctuations during Scenario 2.

### Adverse events monitoring

One participant reported feeling hot during Scenario 2 and the simulation was immediately terminated. The symptom quickly abated, and he completed remaining procedures but his heart rate, cortisol, and PIQ data were excluded from analyses. None of the other participants reported any problems or finding the scenarios too distressing. A few suggested increasing their intensity (e.g., include guns). One participant without PTSD had a nightmare during the week after the VR challenge. No other participants reported increased symptoms or distress. PCL scores before and after the VR challenge were virtually identical. A repeated measures ANOVA found pre-challenge Mean (SD) = 27.41 (11.00), post-challenge Mean (SD) = 27.22 (10.57),  $F(1, 31) = 0.32$ ,  $p = .860$ .

### Immediate involuntary thoughts – suppression group

Thought frequency during the suppression task ranged from 0 ( $n = 1$ ) to 26, Mean (SD) = 11.07 (8.76). There were 38 reports of different types of content: 50% was visual imagery and 18% was sounds (e.g., weaponry). Events and actions (e.g., patrolling) each constituted 11% of content. Emotions (vulnerability) and sensations (holding a rifle) comprised 5% of content. The remaining 5% was non-specific (general thoughts about the scenarios).

Significant zero-order predictors of involuntary thoughts were CAPS hyperarousal symptoms, lower cortisol at the first visit, and higher heart rate at baseline and during both scenarios except for Segment 4 of Scenario 2. Coefficient values ranged from  $r = .547$ ,  $p = .035$  to  $r =$

**Table 2.** Regression Analyses: Immediate, Weekly, and Delayed Involuntary Thoughts.

Involuntary Thoughts Immediately After Scenarios					
Model	Beta	Adjusted R2	Std. Error of the Estimate	R <sup>2</sup> Δ	p
Step 1. Mean heart rate, final segment of Scenario 2	.724	.480	6.192	.524	.005
Step 2. Mean heart rate, final segment of Scenario 2; Cortisol level, Visit 1	.586 −.468	.669	4.942	.201	.007 .022
Involuntary Thoughts During the Week					
Model	Beta	Adjusted R2	Std. Error of the Estimate	R <sup>2</sup> Δ	p
Step 1. SUDS* Scenario 2 Segments 1 & 2	.504	.228	0.656	.254	.004
Step 2. SUDS* Scenario 2 Segments 1 & 2; Experimental Condition**	.547 −.444	.410	0.573	.195	.004
Involuntary Thoughts at One Week Follow Up					
Model	Beta	Adjusted R2	Std. Error of the Estimate	R <sup>2</sup> Δ	p
Step 1. SUDS* Scenario 2 Segments 3 & 4	.435	.160	1.376	.189	.016
Step 2. SUDS* Scenario 2 Segments 3 & 4; Cortisol, Visit 1 Sample	.506 −.479	.370	1.191	.225	.002 .003

\* SUDS = Subjective Units of Distress Scale.

\*\* Experimental Condition variable coded as 1 = Suppression, 2 = Recounting.

−.647,  $p = .009$ . Gender and education were included as covariates in the regression analysis because being female and better educated predicted lower cortisol level at the first visit. Race/ethnicity was included in the model because minorities had higher heart rates at baseline and throughout both scenarios. Results are shown in Table 2. Lower cortisol and higher heart rate during the last segment of Scenario 2 significantly predicted 67% of the variance in thought frequency.

Post hoc analyses were conducted to explicate the relationships amongst predictors. CAPS Hyperarousal symptom subscale scores and cortisol level did not correlate with each other or with heart rate. Of the five CAPS hyperarousal symptoms, startle response was the only significant predictor of involuntary thoughts,  $r = .652$ ,  $p = .008$ ; and also predicted lower cortisol levels at the first visit,  $r = -.362$ ,  $p = .042$ . The correlation between startle responses and involuntary thoughts was not significant after controlling for cortisol level, suggesting that this variable mediated the relationship between startle responses and thought frequency.

### Weekly thought frequency

Twenty-two participants (69%) thought about the scenarios at least once during the week and two made efforts to suppress thoughts. Mean (*SD*) thought frequency ratings were 1.91 (0.73) for the sample, 2.19 (0.54) for the Suppression group, and 1.63 (0.81) for the Recounting group. Experimental condition (thought suppression) predicted higher weekly thought frequency ratings,  $r = -.389$ ,  $p = .028$ . For the Suppression group, the frequency of involuntary thoughts immediately after the VR challenge

predicted higher weekly thought frequency ratings,  $r = .622$ ,  $p = .013$ . For the Recounting group, none of the immediate free recall memory variables predicted frequency ratings. Additional significant zero-order predictors of higher weekly thought frequency ratings included more severe weekly PCL scores, greater immersion in the scenarios, and higher SUDS ratings after Scenario 1 and before and throughout Scenario 2. Coefficient values ranged from  $r = .376$ ,  $p = .034$  to  $r = .504$ ,  $p = .003$ . As shown in Table 2, a regression analysis that controlled for race/ethnicity found that suppression and higher SUDS ratings during the first two segments of Scenario 2 (patrol) predicted higher weekly thought frequency ratings and accounted for 41% of the variance.

### Involuntary thoughts at follow-up

The number of thoughts during the suppression task ranged from 0 ( $n = 3$ ) to 61 ( $n = 1$ ). The mean (*SD*) thought frequency ranking score was 3.42 (1.48) and the mode was 3 (indicating 3–5 thoughts). There were 62 reports of different types of content. Visual imagery was again the most frequent (73%), followed by sounds (11%), non-specific content (11%), and actions (5%).

For the Suppression Group, a greater number of involuntary thoughts immediately after the VR challenge predicted higher thought frequency ranking scores at follow up,  $r = .865$ ,  $p < .001$ . For the Recounting group, none of the immediate free recall memory variables predicted thought frequency scores. Scores for the two groups were similar: Mean = 3.60, *SD* = 1.55 (Suppression); Mean = 3.25, *SD* = 1.44 (Recounting); and experimental condition did not significantly predict involuntary thought scores.

Additional significant zero-order predictors of higher involuntary thought scores were lower levels of cortisol at the first visit, poorer delayed recall of event sequences, and higher SUDS ratings during the first four segments of Scenario 2. Coefficient values ranged from  $r = -.376, p = .037$  to  $r = .435, p = .016$ . Gender and education were included in the regression to control for their effects on cortisol level. As shown in Table 2, lower cortisol level and higher SUDS ratings during the Scenario 2 attack segments significantly predicted 37% of the variance in involuntary thought scores.

### Summary: predictors of involuntary memories

Lower cortisol level at the first visit and faster heart rate during the last segment of Scenario 2 were robust predictors of more frequent involuntary thoughts immediately after the VR challenge and accounted for two-thirds of the variance in thought frequency. Lower cortisol level at the first visit also mediated the relationship between startle responses and more frequent involuntary thoughts immediately after the VR challenge. Higher SUDS ratings for the initial segments of Scenario 2 (patrol) and suppressing thoughts after the VR challenge predicted higher weekly thought frequency ratings and accounted for 41% of the variance in frequency ratings. Higher SUDS ratings for Scenario 2 attack segments and lower cortisol level at the first visit predicted more frequent thoughts during the suppression task at one-week follow-up, accounting for 37% of the variance in involuntary thought scores. Greater immersion in the scenarios, weekly PCL scores, and impaired delayed recall of event sequences were significant zero-order predictors that were excluded from final equations after controlling for variance shared with other variables.

### Immediate free recall – recounting group

Participants recalled a mean (SD) of 6.31 (2.80) danger-related stimuli, 4.59 (2.33) neutral stimuli, 1.75 (1.48) danger-related events, 5.91 (3.33) neutral events, and 3.38 (3.46) event sequences. Better recall of *neutral events* was predicted by less dissociation during scenarios,  $r = -.681, p = .004$ ; and by higher mean heart rate during the first segment of Scenario 2 (patrol),  $r = .514, p = .042$ . A regression analysis that controlled for race/ethnicity identified dissociation as the only significant predictor of memory for neutral events,  $\text{Beta} = -.681, SE = 2.522, p = .004, \text{adjusted } R^2 = .425$ . The only predictor of better recall of *danger-related events* was lower level of education,  $r = -.532, p = .034$ ; which also predicted superior recall of *event sequences*,  $r = -.530, p = .035$ . Impaired recall of event sequences was predicted by greater dissociation during scenarios and lower mean heart rate during Scenario 2 Segments 1, 3, 4 and 5, with coefficient values ranging from  $r = -.498, p = .050$  to  $r = -.668, p = .005$ . A regression analysis that controlled for race/ethnicity

found that dissociation and lower heart rate during the first segment of Scenario 2 predicted poorer recall of event sequences:  $\text{Beta} (\text{dissociation}) = -.540, \text{Beta} (\text{heart rate}) = .413, SE = 2.348, p = .044, \text{adjusted } R^2 = .539$ . Superior delayed recall of NYU paragraphs was the only significant predictor of free recall of *neutral stimuli*,  $r = .521, p = .038$ . Higher NYU Memory Test scores for immediate and delayed recall of paragraphs and paired associates predicted better recall of *danger-related stimuli*, with coefficients ranging from  $r = .555, p = .026$  to  $r = .604, p = .013$ . The regression analysis identified superior delayed recall of NYU paragraphs as the only significant predictor of memory for *danger-related stimuli*,  $\text{Beta} = .604, SE = 2.314, p = .013, \text{adjusted } R^2 = .319$ .

### Delayed free recall and multiple-choice memory test scores

Test scores and regression analyses results are shown in Tables 3 and 4, respectively. A MANCOVA that controlled for Age, Education, Race, and NYU Memory Test scores found no significant effects of Experimental Condition on the overall free recall memory scores was not significant, Pillai's Trace  $F(3, 21) = 1.083, p = .378$ . There were no significant group differences in total Delayed Free Recall Memory scores,  $F(1, 23) = 2.497, p = .128$ ; total Delayed Free Recall Memory – Neutral scores,  $F(1, 23) = 1.713, p = .203$ ; or total Delayed Free Recall Memory – Danger Scores,  $F(1, 23) = 1.721, p = .203$ . Significant zero-order predictors of better recall of *neutral events* included lower SUDS before Scenario 2, and NYU Memory Test scores for delayed recall of paragraphs and immediate and delayed recall of paired associates.

Coefficient values ranged from  $r = -.351, p = .049$  to  $r = .411, p = .019$ . The regression analysis found that NYU Paragraph Delayed Recall scores was the only significant predictor of delayed recall of *neutral events*:  $\text{Beta} = .411, SE = 3.615, \text{adjusted } R^2 = .141, p = .019$ . Significant zero-order predictors of superior recall of *danger-related events* included younger age, lower BDI scores, higher NYU Memory Test scores for delayed recall of paragraphs, and lower SUDS ratings before Scenario 2 and during its first four segments. Coefficient values ranged from  $r = -.366, p = .039$  to  $r = -.613, p < .001$ . The regression analysis found that age and delayed paragraph recall scores predicted approximately 44% of the variance in memory for *danger-related events*.

A partial correlation that controlled for group differences in NYU Memory Test scores for delayed recall of paired associates found that Recounting scenarios improved memory for *event sequences*,  $r = .360, p = .050$ . Other significant zero-order predictors of enhanced recall of *event sequences* included more years of education, higher NYU Memory Test scores for delayed recall of paragraphs and for immediate recall of paired associates, and lower levels of dissociation during scenarios. Coefficient values ranged from  $r = -.352, p = .048$  to  $r = -.381, p = .035$ .

**Table 3.** Delayed Free Recall and Multiple-Choice Test Scores

	Suppression Group N = 16	Recounting Group N = 16 Mean (SD)	Total Sample N = 32
Total Score	19.31 (12.11)	21.47 (12.47)	
Neutral Total	10.72 (6.57)	11.03 (6.15)	
Danger Total	5.75(3.41)	5.88 (3.09)	
Danger-related Events (Raw Score)	1.34 (1.39)	1.47 (1.77)	1.41 (1.57)
Danger-related Events Percentile Score*	2.69 (1.14)	2.69 (0.95)	2.69 (1.03)
Neutral Events	5.25 (3.61)	6.19 (4.24)	5.72 (3.91)
Event Sequences	2.84 (3.20)	4.56 (4.05)	3.70 (3.70)
Danger-related Stimuli	4.41 (2.56)	4.41 (1.69)	4.41 (2.13)
Neutral Stimuli	5.47 (3.70)	4.84 (2.46)	5.16 (3.11)
Multiple Choice Danger Items % Correct	36.27 (12.57)	37.90 (14.52)	37.09 (13.38)
Multiple Choice Neutral Items % Correct	39.15 (12.67)	42.43 (8.78)	40.79 (10.85)
Pillai's Trace $F(3, 21) = 1.083, p = .378$			
Total Score $F(1, 23) = 2.497, p = .128$			
Total Neutral $F(1, 23) = 1.713, p = .203$			
Total Danger $F(1, 23) = 1.721, p = .203$			

\* Danger Narrative Events Memory Classification score coded as 1 = < 25th percentile; 2 = < 50th percentile; 3 = < 75th percentile; 4 < 100th percentile.

**Table 4.** Regression Analyses, Delayed Free Recall and Multiple-Choice Memory Test Scores

Dependent Variable: Danger-related Events					
Model	Beta	Adjusted R2	Std. Error of the Estimate	R <sup>2</sup> Δ	p
Step 1. Age	-.584	.318	0.825	.341	.001
Step 2. Age; NYU Paragraph Delayed Recall	-.512 .374	.438	0.749	.135	.012
Dependent Variable: Neutral Stimuli					
Model	Beta	Adjusted R2	Std. Error of the Estimate	R <sup>2</sup> Δ	p
Step 1. HRV*	.391	.122	2.659	.153	.033
Step 2. HRV* Pre Scenario 2 SUDS*	.462 -.432	.285	2.400	.182	.011
Dependent Variable: Multiple Choice Neutral Items					
Model	Beta	Adjusted R2	Std. Error of the Estimate	R <sup>2</sup> Δ	p
Step 1. Cortisol, Post-scenarios	.592	.326	8.099	.350	.001
Step 2. Cortisol, Post-scenarios; Scenario 1 Heart Rate Δ	.490 -.431	.489	7.053	.175	.005
Step 3. Cortisol, Post-scenarios; Scenario 1 Heart Rate Δ NYU Paragraph Delayed Recall	.452 -.375 .315	.574	6.442	.094	.020

\* HRV = Baseline Heart Rate Variability; SUDS = Subjective Units of Distress Scale.

A regression analysis that controlled for group differences in NYU Memory Test scores for delayed recall of paired associates and for race/ethnicity identified dissociation as the only significant predictor of memory for *event sequences*: Beta =  $-.381$ ,  $SE = 3.476$ , adjusted  $R^2 = .115$ ,  $p = .035$ .

Higher SUDS ratings before Scenario 2 predicted poorer recall of *neutral stimuli*,  $r = -.394$ ,  $p = .026$ , whereas higher baseline heart rate variability predicted better recall,  $r = .391$ ,  $p = .033$ . As shown in Table 4, the regression analysis found that these two variables predicted 29% of the variance in memory for *neutral stimuli*. No significant predictors of delayed free recall memory of *danger-related stimuli* and for *multiple choice test danger-related items*

were identified for the sample as a whole. For the Recounting Group, better immediate free recall predicted better delayed free recall of all variables except for delayed recall of *neutral events*. Coefficient values ranged from  $r = .507$ ,  $p = .045$  to  $r = .769$ ,  $p < .001$ . Immediate recall of danger-related events was the only predictor of *danger-related multiple-choice scores*,  $r = .584$ ,  $p = .018$ .

Significant zero-order predictors of higher scores for *multiple choice test neutral items* were male gender, higher NYU Memory Test immediate and delayed paragraph recall scores. Higher cortisol level after the scenarios and greater increase in heart rate during Scenario 1 predicted lower scores. Coefficient values ranged from  $r = .372$ ,  $p = .036$  to  $r = .592$ ,  $p < .001$ . Participants in the

Recounting group also recalled a greater percentage of *multiple choice test neutral items*,  $r = .422, p = .023$ , controlling for gender, race/ethnicity, and NYU Memory Test scores for delayed recall of paired associates. As shown in [Table 4](#), a regression analysis that controlled for gender, race/ethnicity and NYU Memory Test scores found that higher cortisol level after the scenarios, Scenario 1 heart rate change scores, and NYU Memory Test scores predicted 57% of the variance in scores for *multiple choice test neutral items*.

### Summary: predictors of voluntary memory

Verbal memory ability, dissociation, and physiological arousal were the strongest predictors of voluntary memory. Superior delayed recall of paragraphs predicted better immediate recall of stimuli in the scenarios, better delayed recall of events in the scenarios, and higher multiple-choice test scores for neutral items. Dissociating during scenarios impaired immediate recall of neutral events and delayed recall of event sequences. Dissociation and lower heart rate during the first segment of Scenario 2 also impaired immediate recall of event sequences, accounting for 54% of the variance in scores. Higher baseline heart rate variability predicted better delayed free recall of neutral stimuli, whereas higher SUDS ratings before the second scenario impeded recall of neutral stimuli. Delayed multiple choice recognition memory for neutral items was enhanced by higher cortisol level after the scenarios but impaired by greater increase in heart rate from baseline during Scenario 1, which together predicted 57% of the variance in scores. Younger age and verbal memory ability significantly predicted better delayed recall of dangerous events and accounted for 44% of the variance in scores. For the Recounting Group, immediate recall of dangerous events predicted better delayed free recall of danger-related stimuli and delayed

multiple-choice recognition memory for danger-related items (see [Figure 2](#) for a tabular view of main predictors/outcomes).

## Discussion

### Summary of the results

This study utilised a novel, prospective VR research design to model the development of voluntary and involuntary episodic memories of stressful events. To our knowledge, it is one of the few if not the first prospective study of emotional episodic memories that included multivariate predictors sampled across multiple response and time domains and fine-grained analyses of memory characterised according to content and valence. Such analyses are required to test etiological theories of PTSD re-experiencing symptoms and trauma memories. In support of the “dual-process” trauma memory theory (Brewin, 2011, 2014; Brewin et al., 2010; Kaye et al., 2019) dissociation, lowered heart rate, and higher distress during scenarios selectively impaired episodic memory features (neutral stimuli, events, and event sequences), with no effect on memory for danger-related elements. Suppressing thoughts of the scenarios (compared to recounting) also selectively impaired delayed recall of event sequences and delayed recognition of neutral multiple-choice items, with no effect on danger-related memory elements.

Physiological arousal also selectively affected memory for neutral events and stimuli, but only a few significant effects were identified. The direction of effects also varied according to the timing of the arousal and whether testing involved free recall or recognition. Greater arousal, as indexed by faster heart rate during the first segment of Scenario 2 (patrol), improved immediate free recall of event sequences; and higher levels of cortisol

Outcome Predictor	Impaired episodic memory features	Impaired delayed recall of event sequences	Delayed recognition of neutral multiple-choice items	Danger-related memory elements	Memory for neutral events/stimuli	Involuntary memories frequency
Dissociation	*					
Lowered HR	*					
Higher distress during VR scenarios	*					
Suppressing thoughts		*	*	No effect		*
Physiological arousal					*	*
Emotional stress response						*
Cortisol at baseline						*

Note: \* represents significant association between the predictor and the corresponding outcome.

**Figure 2.** Tabular view of main predictors/findings. Note: \* represents significant association between the predictor and the corresponding outcome.

after the scenarios enhanced delayed verbal recognition memory for neutral multiple-choice test items. However, higher levels of arousal before the scenarios (i.e., lower heart rate variability at baseline), and greater increase in heart rate from baseline during the first scenario predicted lower scores for neutral multiple-choice test items. Clearly more research with repeated tonic and phasic sampling of multiple indices of stress responses is needed. However, the results of this study suggest a kind of primacy-recency effect in which higher tonic physiological arousal and greater physiological reactivity during the first few minutes of exposure to stressful events impair episodic memory and higher arousal towards the end of exposure enhances it.

In the present study, a higher frequency of involuntary memories of the scenarios was predicted by physiological and emotional stress responses and to a lesser extent, suppression. The effects of physiological reactions to the scenarios and thought suppression were short-lived. Higher heart rate during the scenarios predicted more frequent involuntary thoughts immediately after the scenarios, but not at one-week follow up; and thought suppression was only associated with higher subjective ratings of weekly thought frequency. More frequent involuntary thoughts immediately after the scenarios did predict more frequent thoughts during the suppression task one week later, suggesting a mechanism by which physiological reactions during exposure to stressors could precipitate persistent involuntary memories.

General thought suppression ability (Stroop scores) did not predict the frequency of involuntary thoughts. The Stroop test might not have been challenging enough to detect cognitive control deficits that might be observed on more demanding tasks. Involuntary thoughts are harder to suppress when they elicit more intense emotions and when there are competing cognitive demands (Nixon et al., 2007, 2009). Thus, it is possible that the relationship between suppression ability and involuntary thoughts is mediated by the amount of distress evoked by the memories and/or the inability to suppress them. Conversely, it is also possible that there is no relationship between inhibitory control and involuntary thoughts, as some recent studies suggest (Barzykowski, Radel, et al., 2019). In our study, suppression coupled with higher distress during the first two segments of Scenario 2 (patrolling) predicted higher weekly thought frequency ratings. Higher distress during Scenario 2 attack segments, along with lower cortisol level at the first visit, also predicted more frequent involuntary memories at follow up. We interpreted distress during the patrol segments as reflecting anticipatory anxiety and distress during the attack as emotional threat responses. The results suggest that anticipatory anxiety and post-exposure thought suppression temporarily increase the perceived frequency of involuntary thoughts, whereas emotional threat responses evoked by perceived danger may promote more persistent problems inhibiting stressful memories.

Lower level of cortisol at the initial study visit was the only variable that predicted the frequency of involuntary thoughts initially and at follow up. This variable and heart rate during the final segment of the second scenario accounted for two-thirds of the variance in the frequency of involuntary thoughts immediately after the scenarios. Lower level of cortisol at the initial study visit and distress level during Scenario 2 attack segments also predicted over one-third of the variance in involuntary thought scores at follow up. Cortisol level before and after the scenarios did not predict the frequency of involuntary thoughts, and so the effects did not appear to be directly related to an acute stress response. We speculate that lower cortisol at the first visit reflected tonic hyperarousal because this variable mediated the relationship between the severity of monthly startle responses and the initial frequency of involuntary thoughts. The finding is intriguing because it suggests a mechanism by which pre-existing hyperarousal symptoms could increase the risk of PTSD. Startle responses are a feature of several anxiety disorders (McTeague et al., 2010) and a history of anxiety disorders is a risk factor for PTSD (Ozer et al., 2003). PTSD has also been associated with the tendency to over-generalize startle responses across contexts regardless of the probability of exposure to aversive stimuli (Grillon, 2008). Grillon (2008) characterised context-insensitive startle responses as anxiety-potentiated and as arising in situations in which exposure to adverse events is anticipated but not predictable. Elsewhere we have suggested that this type of generalised anxiety could increase PTSD risk or impede its natural remission (Malta, 2012). Context-insensitive startle responses have also been observed in homozygous carriers of the short allele of the serotonin transporter gene-linked polymorphic region (5-HTTLPR) (Karl, Malta et al., unpublished) and this allele has also been linked to altered functioning in neural emotion regulation circuits (Canli & Lesch, 2007; Heinz et al., 2005). The results of the present study suggest that startle responses could be a biomarker of emotional reactivity that increases the PTSD risk in part by promoting the development of involuntary trauma memories.

#### **Link to prior findings:**

These results replicated and augmented those of prior thought suppression research (Wegner et al., 1996), although without a control group that neither suppressed nor recounted memories, it is not possible to know whether the effects were due to active suppression *per se*. The effects of suppression were no longer significant after accounting for those of other variables. The selectivity and relative weakness of the suppression effects is also consistent with previous research (Abramowitz et al., 2001; Wegner et al., 1996) and suggest that posttraumatic amnesia may be more related to impaired encoding during exposure rather than post-exposure memory suppression. Future research should test whether persistent suppression

of memories for longer periods of time leads to more amnesia, as one recent study found (Hulbert et al., 2016).

This study is also one of the few to have assessed involuntary and voluntary memories of the same events. Visual imagery was the most frequent type of involuntary memory content, which is consistent with descriptions of trauma-related involuntary memories (Brewin et al., 2010). The largely perceptual nature of involuntary memories was also in accordance with the dual-process trauma memory model (Brewin, 2011, 2014; Brewin et al., 2010) and with the finding that superior visual memory prospectively predicts the development of war-related PTSD (Marx et al., 2009). We did not find any association between the ability to recall danger-related elements and the frequency of involuntary memories, but this may have been due to our not having included visual-perceptual tests of memories for the scenarios. The study also found no evidence of a direct relationship between voluntary and involuntary memories, but they were indirectly related through the shared predictors of physiological arousal, emotional distress, and suppression. Higher physiological arousal selectively impaired episodic memory. Greater distress during scenarios impaired voluntary memory for neutral elements but increased involuntary memories during the week and at follow up. The results were thus consistent with the dual-process trauma memory model (Brewin, 2011, 2014; Brewin et al., 2010).

#### **Counterintuitive results:**

Contrary to expectation, neither physiological arousal nor distress strengthened danger-related memories. The VR scenarios evoked moderate stressor responses, and more intense reactions may have been needed to produce effects. The study identified few predictors of danger-related memories, other than immediate recall memory, age, and verbal memory ability. Younger age and better delayed recall of paragraphs emerged as strong predictors of danger-related events, jointly accounting for 44% of the variance in delayed free recall of danger-related events scores. Younger age is a PTSD risk factor (Norris et al., 2002), and future research should investigate whether enhanced recall of adverse events is a mechanism by which this risk factor operates. Better scores on the NYU Memory Test (Kluger et al., 1999) was the most consistent predictor of voluntary recall. Superior performance on this test predicts larger hippocampal volumes (Golomb et al., 1996), and smaller hippocampal volumes and deficient verbal associative memory have been associated with PTSD (Guez et al., 2011; Karl et al., 2006), though mechanism by which these risk factors operate are unknown.

#### **Novelty:**

One of the novel features of the study was the use of VR technology to create laboratory stimuli capable of producing valid stressor responses without creating undue levels of distress.

The VR challenge paradigm evoked mild-to-moderate stressor responses and intrusive memories and appeared to be reasonably well-tolerated. A strength of the study was the use of a symptomatic trauma-exposed sample, but this introduced the risk of symptom exacerbation. The only instance of this that was reported was the nightmare experienced by one participant. We do not know whether this was elicited by the VR challenge or other procedures, or how this compares to other PTSD symptom challenge studies (Patel et al., 2012), which do not generally include information about symptom exacerbation. It would have been preferable if independent evaluators conducted the debriefings. However, participants appeared to offer honest feedback, which in all cases were suggestions for increasing the intensity of the simulations. Despite these encouraging findings, we would suggest some caution conducting VR research with samples that have more severe symptoms than our sample did. We would also recommend that researchers who plan to conduct VR research with clinical samples inform participants of the risks of transient symptom increase and simulator sickness and monitor for adverse reactions before and after participation, as we did.

#### **Ethical aspects**

This study utilised a novel VR challenge paradigm to study emotional memories. The need to use valid stimuli of course must be balanced with ensuring the welfare of subjects, and we tried to create VR scenarios that would be evocative but not overly distressing. We assessed the level of distress during the study, at the final visit one week after the VR challenge, and by telephone one week after the study was completed. We found no significant changes in PTSD symptoms assessed with the PCL at visit 1 vs. visit 3. Other than the nightmare reported by one participant, there were no reports of increased PTSD symptoms during or after the study. We do not know whether the nightmare was due to the VR challenge or to other procedures or factors. We also do not know whether this also happens with other symptom challenge paradigms, which have not reported on any symptoms changes after the study is completed (Lanius et al., 2010; Liberzon et al., 1999; Pitman et al., 1987; Shipherd & Beck, 1999). We also assessed for any distress one week after the study was completed, and again received no report of any problems. It would have been preferable to have study debriefings and follow up calls administered by independent evaluators. However, the participants appeared to offer honest feedback throughout the study, which in all cases was suggestions for *increasing* the intensity of the scenarios. It was not practical for us to assess for any distress or problems beyond the two-week period, but participants with PTSD were provided with referrals and all participants were encouraged to contact us if they felt the need for assistance. We would suggest that researchers take care when using symptoms challenge paradigms, especially with

clinical samples with more severe PTSD, as our sample was composed of Veterans whose symptoms were only mild-to-moderate. Researchers who plan to use VR paradigms with trauma survivors should also inform participants during the consenting process of the possibility of transient increases in PTSD symptoms and simulator sickness, as we did in the present study.

### *Recap of the main findings*

To summarise the main findings and their implications, the results suggest that posttraumatic amnesia is most strongly related to deficient verbal memory ability, dissociative reactions to stressors, and to a lesser extent, physiological arousal. Dissociation and physiological arousal selectively affected neutral episodic memory elements, and greater distress during scenarios impaired voluntary memory but increased the frequency of involuntary memories. The collective findings provide evidence in support of the hypothesis that disparate memory systems underlie the formation of involuntary and voluntary trauma memories (Brewin, 2014).

The results also support a hypothetical model of involuntary memory formation in which tonic (low cortisol) and acute (heart rate) physiological stress responses promote the development of involuntary memories immediately after exposure to stressors. In turn, the initial tendency to develop involuntary memories predicts more frequent involuntary memories at follow-up. Tonic physiological stress responses (low cortisol) and higher emotional threat responses during exposure maintain the tendency to have difficulty suppressing memories when they are reactivated a week after exposure. Post-exposure thought suppression and apprehension during stressor exposure increase the perception of having more frequent involuntary memories during the week after exposure, which could compound the distress experienced after trauma exposure. The collective results suggest that pre-exposure physiological hyperarousal, memory ability, and emotional and physiological reactions during stressor exposure influence the initial development of intrusive memories and other PTSD re-experiencing symptoms, whereas suppression and avoidance might play a greater role in maintaining the symptoms. The findings also suggest that PTSD screening programmes for Veterans, first responders, and disaster survivors might be enhanced by including assessments of mental health history and physiological stress responses to improve identification of those at risk to develop PTSD.

### *Limitations of the study*

Despite the promising results of this study, its complexity comes with a series of limitations, briefly summarised below:

#### (a) Sample issues and the absence of a control group

To minimise demand characteristics, we did not control for the extent to which participants talked about scenarios

during the week after exposure, and therefore we do not know whether the results were influenced by the behaviours of participants during the week. Because this study introduced a new type of laboratory stressor, we recruited a minimally adequate number of participants and did not include a control group that neither suppressed nor recounted memories. Exclusions and attrition led to a final sample that was smaller than anticipated and limited statistical power. There were also too few females to conduct meaningful analyses of gender effects. The study's comprehensive approach to modelling predictors of emotional memory necessitated numerous analyses. Controlling for inflated experiment-wise Type 1 error would have resulted in overly stringent alpha levels, further compromising statistical power and increasing the probability of Type II errors. Instead, our interpretations of results emphasized the effects of predictors that accounted for meaningful amounts of variance in the outcome variables and findings that were consistent with previous research.

#### (b) Inducing immersion

We had participants remove the HMD in between the VR scenarios, which could have negatively affected their level of immersion. This, in turn, could have affected the vividness of memories, making them more intrusive and easier to recall. However, the initial analyses we conducted to identify significant zero order correlations (described under Data Analyses) did not find that level of immersion was significantly correlated with any of the involuntary or voluntary memory dependent variables. Nonetheless, minimal disruptions to immersion are unavoidable and VR researchers should try to find ways to maximise immersion in virtual environments.

#### (c) Lack of analyses of personality traits

Research suggests that personality traits modulate the neural mechanisms underlying the emotional episodic memory (Dolcos et al., 2017). We did not collect personality traits data. Further studies should examine the role of personality traits on the complex relationships unveiled in our study.

#### (d) Measuring immersion:

We chose to use an adapted instrument, the Immersion/Presence Questionnaire, as a process measure during the VR challenge, because of its brevity and ease of administration during the VR challenge. We had to ensure that all procedures were completed within a time window in order to standardise the amount of time between the pre- and post-challenge cortisol assessments. However, no psychometric studies were conducted with this instrument. The internal validity of the study would have been

strengthened by the use of an instrument with known psychometric properties.

(e) Cortisol:

Variability in the amount of time elapsed between cortisol samples before and after the challenge could have affected the findings. Because cortisol takes approximately 15 min. to manifest in the saliva, 25-30 min. to peak, and 60 min. to decline (Kirschbaum et al., 1993), it is possible that post-challenge cortisol samples taken too early (20 min.) or too late (40 min.) would not have been at peak. However, as we described earlier (see footnote 2), there were no significant correlations between the duration of time between pre- and post-challenge cortisol samples and of the involuntary or delayed voluntary memory dependent variables. We did find a significant correlation with the Immediate Free Recall of Danger stimuli. However, including the amount of time between cortisol samples as a control variable did not change the results of the regression analysis. Therefore, although it is important to minimise variability in cortisol sampling times, we did not identify any significant effects on our study results. Future VR researchers should develop paradigms that would allow for frequent, repeated sampling of cortisol throughout exposure to VR scenarios, as it would provide more precise data on stress hormone responses.

(f) Time to recount:

We did not record the amount of time it took for subjects to recount memories of the scenarios and therefore could not control for any potential effects of this on analyses.

(g) No recounting group:

As noted, we also did not have a control group that neither recounted memory nor actively suppressed thoughts of the scenarios.

(h) The construct "Dissociation":

The construct "dissociation" may have been poorly operationalised by the use of a 3-item score from the IPQ, which is another limitation that should be addressed in future studies.

## Conclusion

Despite these caveats, the study provided novel data on involuntary and voluntary memories of stressful events and hypotheses for testing in future research. The study also contributes an ecologically valid and well-tolerated VR challenge paradigm that can be utilised to advance our understanding of emotional episodic memory processes and their role in the development of stress-related

disorders. Clinical scientists should continue to explore ways to utilise VR technology to conduct research on the effects of psychological stress on emotional and physical well-being.

## Notes

1. The study was also reviewed and approved for funding by the National Institutes of Health loan repayment grant program, as well as the Weill Medical College of Cornell University's research committee that awarded internal departmental funding.
2. We used this method to normalize variables because using z-score transformations did not sufficiently normalize kurtotic distributions. We also did not want to further reduce the small sample size by excluding outliers.
3. Two subjects completed the diagnostic assessment and the VR challenge within the same week due to scheduling issues.
4. Two subjects' visits commenced at 5:00 and 5:15 PM due to their tardiness. For the majority, visits were scheduled within one hour (77%-90%) or two hours (91-97%) of each other.
5. We conducted a series of analyses to identify potential effects of the variability in the sampling time. A one-way ANOVA (with the outlier removed) found a longer average duration in Suppression Group compared to the Recounting Group,  $F(1,29) = 4.958, p = .034$ . However, the duration of time between cortisol samples was not significantly correlated with the post-VR challenge level of cortisol or with the pre-post challenge change in cortisol level. Nor was it correlated with the frequency of involuntary thoughts about the scenarios or making attempts to suppress thoughts during the week. It was also not correlated with any of the delayed recall and recognition memory variables. There was one significant correlation with the Immediate Free Recall of Danger stimuli. However, the regression analysis in the Results section was repeated, controlling for this variable, and the results were unchanged.

## Acknowledgements

The authors wish to thank the study research assistants, Amy Elitzer and Ashley Okamoto, for their invaluable work with the scoring of memory tests.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This study was supported by NIH Loan Repayment and Weill Medical College of Cornell University Faculty Grants awarded to the first author. These agencies had no involvement in the design or execution of the research or in the writing of research manuscripts.

## ORCID

Cezar Giosan  <http://orcid.org/0000-0002-1260-6830>  
Albert A. Rizzo  <http://orcid.org/0000-0003-2647-7731>

## References

- Abramowitz, J. S., Tolin, D. F., & Street, G. P. (2001). Paradoxical effects of thought suppression: A meta-analysis of controlled studies.

- Clinical Psychology Review*, 21(5), 683–703. doi:10.1016/S0272-7358(00)00057-X
- Admon, R., Milad, M. R., & Hendler, T. (2013). A causal model of post-traumatic stress disorder: Disentangling predisposed from acquired neural abnormalities. *Trends in Cognitive Sciences*, 17(7), 337–347. <https://doi.org/10.1016/j.tics.2013.05.005>
- Althouse, A. D. (2016). Adjust for multiple Comparisons? It's Not that Simple. *The Annals of Thoracic Surgery*, 101(5), 1644–1645. <https://doi.org/10.1016/j.athoracsur.2015.11.024>
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders 4th Edition*.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders 5th Edition*.
- Anderson, M. C., & Levy, B. J. (2009). Suppressing Unwanted memories. *Current Directions in Psychological Science*, 18(4), 189–194. <https://doi.org/10.1111/j.1467-8721.2009.01634.x>
- Barkley, R. A., & Murphy, K. R. (2005). *Attention-Deficit Hyperactivity disorder, third Edition: A clinical Workbook (3rd edition)*. The Guilford Press.
- Barzykowski, K., & Niedźwieńska, A. (2016). The effects of instruction on the frequency and characteristics of involuntary autobiographical memories. *PLOS ONE*, 11(6), Article e0157121. <https://doi.org/10.1371/journal.pone.0157121>
- Barzykowski, K., Niedźwieńska, A., & Mazzoni, G. (2019). How intention to retrieve a memory and expectation that a memory will come to mind influence the retrieval of autobiographical memories. *Consciousness and Cognition*, 72, 31–48. <https://doi.org/10.1016/j.concog.2019.03.011>
- Barzykowski, K., Radel, R., Niedźwieńska, A., & Kvavilashvili, L. (2019). Why are we not flooded by involuntary thoughts about the past and future? Testing the cognitive inhibition dependency hypothesis. *Psychological Research*, 83(4), 666–683. <https://doi.org/10.1007/s00426-018-1120-6>
- Barzykowski, K., & Staugaard, S. R. (2018). How intention and monitoring your thoughts influence characteristics of autobiographical memories. *British Journal of Psychology*, 109(2), 321–340. <https://doi.org/10.1111/bjop.12259>
- Beck, J. G., Gudmundsdottir, B., Palyo, S. A., Miller, L. M., & Grant, D. M. (2006). Rebound effects following deliberate thought suppression: Does PTSD make a difference? *Behavior Therapy*, 37(2), 170–180. <https://doi.org/10.1016/j.beth.2005.11.002>
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck depression Inventory-II*. Psychological Corporation.
- Bergouignan, L., Nyberg, L., & Ehrsson, H. H. (2014). Out-of-body-induced hippocampal amnesia. *Proceedings of the National Academy of Sciences*, 111(12), 4421–4426. <https://doi.org/10.1073/pnas.1318801111>
- Berntsen, D., & Rubin, D. C. (2014). Involuntary memories and dissociative amnesia: Assessing key assumptions in PTSD research. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, 2(2), 174–186. <https://doi.org/10.1177/2167702613496241>
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Charney, D. S., & Keane, T. M. (1998). *Clinician-administered PTSD scale for DSM-IV*. National Center for Posttraumatic Stress Disorder, Behavioral Science Division.
- Brewin, C. R. (2011). The nature and significance of memory disturbance in posttraumatic stress disorder. *Annual Review of Clinical Psychology*, 7(1), 203–227. <https://doi.org/10.1146/annurev-clinpsy-032210-104544>
- Brewin, C. R. (2014). Episodic memory, perceptual memory, and their interaction: Foundations for a theory of posttraumatic stress disorder. *Psychological Bulletin*, 140(1), 69–97. <https://doi.org/10.1037/a0033722>
- Brewin, C. R., Gregory, J. D., Lipton, M., & Burgess, N. (2010). Intrusive images in psychological disorders: Characteristics, neural mechanisms, and treatment implications. *Psychological Review*, 117(1), 210–232. <https://doi.org/10.1037/a0018113>
- Brewin, C. R., & Smart, L. (2005). Working memory capacity and suppression of intrusive thoughts. *Journal of Behavior Therapy and Experimental Psychiatry*, 36(1), 61–68. <https://doi.org/10.1016/j.jbtep.2004.11.006>
- Canli, T., & Lesch, K.-P. (2007). Long story short: The serotonin transporter in emotion regulation and social cognition. *Nature Neuroscience*, 10(9), 1103–1109. <https://doi.org/10.1038/nn1964>
- Catarino, A., Küpper, C. S., Werner-Seidler, A., Dalgleish, T., & Anderson, M. C. (2015). Failing to forget inhibitory-control deficits compromise memory suppression in posttraumatic stress disorder. *Psychological Science*, 26(5), 604–616. <https://doi.org/10.1177/0956797615569889>
- Dolcos, F., Katsumi, Y., Weymar, M., Moore, M., Tsukiura, T., & Dolcos, S. (2017). Emerging directions in emotional episodic memory. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01867>
- Feise, R. J. (2002). Do multiple outcome measures require *p*-value adjustment? *BMC Medical Research Methodology*, 2(1), 8. <https://doi.org/10.1186/1471-2288-2-8>
- First, M. B., Spitzer, R. L., Williams, J. B. W., & Gibbon, M. (1997). *Structured clinical interview for DSM-IV Axis I disorders*. American Psychiatric Association.
- Gehrman, P., Seelig, A. D., Jacobson, I. G., Boyko, E. J., Hooper, T. I., Gackstetter, G. D., Ulmer, C. S., & Smith, T. C. (2013). Predeployment sleep duration and insomnia symptoms as risk factors for new-onset mental health disorders following military deployment. *Sleep*, 36(7), 1009–1018. <https://doi.org/10.5665/sleep.2798>
- Giosan, C., Malta, L., Jayasinghe, N., Spielman, L., & Difede, J. (2009). Relationships between memory inconsistency for traumatic events following 9/11 and PTSD in disaster restoration workers. *Journal of Anxiety Disorders*, 23(4), 557–561. <https://doi.org/10.1016/j.janxdis.2008.11.004>
- Golden, C., & Freshwater, S. (2002). *The stroop color and word test: A manual for clinical and experimental uses*. Catalog No. 30150M. Stoelting Co.
- Golomb, J., Kluger, A., de Leon, M. J., Ferris, S. H., Mittelman, M., Cohen, J., & George, A. E. (1996). Hippocampal formation size predicts declining memory performance in normal aging. *Neurology*, 47(3), 810–813. doi:10.1212/WNL.47.3.810
- Grillon, C. (2008). Models and mechanisms of anxiety: Evidence from startle studies. *Psychopharmacology*, 199(3), 421–437. <https://doi.org/10.1007/s00213-007-1019-1>
- Guez, J., Naveh-Benjamin, M., Yankovsky, Y., Cohen, J., Shiber, A., & Shalev, H. (2011). Traumatic stress is linked to a deficit in associative episodic memory. *Journal of Traumatic Stress*, 24(3), 260–267. <https://doi.org/10.1002/jts.20635>
- Heinz, A., Braus, D. F., Smolka, M. N., Wrase, J., Puls, I., Hermann, D., Klein, S., Grüsser, S. M., Flor, H., Schumann, G., Mann, K., & Büchel, C. (2005). Amygdala-prefrontal coupling depends on a genetic variation of the serotonin transporter. *Nature Neuroscience*, 8(1), 20–21. <https://doi.org/10.1038/nn1366>
- Holmes, E. A., & Bourne, C. (2008). Inducing and modulating intrusive emotional memories: A review of the trauma film paradigm. *Acta Psychologica*, 127(3), 553–566. <https://doi.org/10.1016/j.actpsy.2007.11.002>
- Hulbert, J. C., Henson, R. N., & Anderson, M. C. (2016). Inducing amnesia through systemic suppression. *Nature Communications*, 7(1), 11003. <https://doi.org/10.1038/ncomms11003>
- Karl, A., Schaefer, M., Malta, L. S., Dörfel, D., Rohleder, N., & Werner, A. (2006). A meta-analysis of structural brain abnormalities in PTSD. *Neuroscience and Biobehavioral Reviews*, 30(7), 1004–1031. <https://doi.org/10.1016/j.neubiorev.2006.03.004>
- Kaye, A. P., Kwan, A. C., Ressler, K. J., & Krystal, J. H. (2019). A computational model for learning from repeated trauma. *BioRxiv*, 659425. <https://doi.org/10.1101/659425>
- Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'trier Social stress test'—A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28(1–2), 76–81. <https://doi.org/10.1159/00019004>

- Kluger, A., Ferris, S. H., Golomb, J., Mittelman, M. S., & Reisberg, B. (1999). Neuropsychological prediction of decline to dementia in nondemented elderly. *Journal of Geriatric Psychiatry and Neurology*, 12(4), 168–179. doi:10.1177/089198879901200402
- Lanius, R. A., Vermetten, E., Loewenstein, R. J., Brand, B., Schmahl, C., Bremner, J. D., & Spiegel, D. (2010). Emotion modulation in PTSD: Clinical and neurobiological evidence for a dissociative subtype. *American Journal of Psychiatry*, 167(6), 640–647. https://doi.org/10.1176/appi.ajp.2009.09081168
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23(1), 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155
- Levy, B. J., & Anderson, M. C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica*, 127(3), 623–635. https://doi.org/10.1016/j.actpsy.2007.12.004
- Liberzon, I., Abelson, J. L., Flagel, S. B., Raz, J., & Young, E. A. (1999). Neuroendocrine and psychophysiological responses in PTSD: A symptom provocation study. *Neuropsychopharmacology*, 21(1), 40–50. doi:10.1016/S0893-133X(98)00128-6
- Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal*, 17(3), 354–381. doi:10.1093/oxfordjournals.eurheartj.a014868
- Malta, L. S. (2012). Allostasis: The emperor of All (trauma-related) maladies. *Clinical Psychology: Science and Practice*, 19(3), 241–259. https://doi.org/10.1111/cpsp.12003
- Marx, B. P., Doron-Lamarca, S., Proctor, S. P., & Vasterling, J. J. (2009). The influence of pre-deployment neurocognitive functioning on post-deployment PTSD symptom outcomes among Iraq-deployed Army soldiers. *Journal of the International Neuropsychological Society: JINS*, 15(6), 840–852. https://doi.org/10.1017/S1355617709990488
- McMillen, J. C., North, C. S., & Smith, E. M. (2000). What Parts of PTSD Are normal: Intrusion, avoidance, or arousal? Data from the Northridge, California, Earthquake. *Journal of Traumatic Stress*, 13(1), 57–75. https://doi.org/10.1023/A:1007768830246
- McTeague, L. M., Lang, P. J., Laplante, M.-C., Cuthbert, B. N., Shumen, J. R., & Bradley, M. M. (2010). Aversive imagery in posttraumatic stress disorder: Trauma recurrence, comorbidity, and physiological reactivity. *Biological Psychiatry*, 67(4), 346–356. https://doi.org/10.1016/j.biopsych.2009.08.023
- Nixon, R. D. V., Cain, N., Nehmy, T., & Seymour, M. (2009). The influence of thought suppression and cognitive load on intrusions and memory processes following an analogue stressor. *Behaviour Therapy*, 40(4), 368–379. https://doi.org/10.1016/j.beth.2008.10.004
- Nixon, R. D. V., Nehmy, T., & Seymour, M. (2007). The effect of cognitive load and hyperarousal on negative intrusive memories. *Behaviour Research and Therapy*, 45(11), 2652–2663. https://doi.org/10.1016/j.brat.2007.06.010
- Norris, F. H., Friedman, M. J., Watson, P. J., Byrne, C. M., Diaz, E., & Kaniasty, K. (2002). 60,000 disaster victims speak: Part I. An empirical review of the empirical literature, 1981–2001. *Psychiatry: Interpersonal and Biological Processes*, 65(3), 207–239. doi:10.1521/psyc.65.3.207.20173
- Ozer, E. J., Best, S. R., Lipsey, T. L., & Weiss, D. S. (2003). Predictors of posttraumatic stress disorder and symptoms in adults: A meta-analysis. *Psychological Bulletin*, 129(1), 52–73. doi:10.1037/0033-2909.129.1.52
- Patel, R., Spreng, R. N., Shin, L. M., & Girard, T. A. (2012). Neurocircuitry models of posttraumatic stress disorder and beyond: A meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 36(9), 2130–2142. https://doi.org/10.1016/j.neubiorev.2012.06.003
- Pause, B. M., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on episodic-like and episodic memory. *Frontiers in Behavioral Neuroscience*, 7, 33. https://doi.org/10.3389/fnbeh.2013.00033
- Pitman, R. K., Orr, S. P., Foa, D. F., de Jong, J. B., & Claiborn, J. M. (1987). Psychophysiological assessment of posttraumatic stress disorder imagery in Vietnam combat veterans. *Archives of General Psychiatry*, 44(11), 970–975. doi:10.1001/archpsyc.1987.01800230050009
- Rassin, E., Merckelbach, H., & Muris, P. (1997). Effects of thought suppression on episodic memory. *Behaviour Research and Therapy*, 35(11), 1035–1038. https://doi.org/10.1016/S0005-7967(97)00059-4
- Regan, C. (1995). An investigation into nausea and other side-effects of head-coupled immersive virtual reality. *Virtual Reality*, 1(1), 17–31. https://doi.org/10.1007/BF02009710
- Rizzo, A. A., Rothbaum, B. O., & Graap, K. (2006). Virtual reality applications for combat-related posttraumatic stress disorder. In C. R. Figley, & W. P. Nash (Eds.), *Combat stress injury: Theory, research, and management* (1st ed., pp. 183–204). Routledge.
- Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology (Cambridge, Mass.)*, 1(1), 43–46. https://doi.org/10.1097/00001648-199001000-00010
- Sachschar, J., Woodward, E., Wichelmann, J. M., Haag, K., & Ehlers, A. (2019). Differential effects of poor recall and memory disjointedness on trauma symptoms. *Clinical Psychological Science*, 7(5), 1032–1041. https://doi.org/10.1177/2167702619847195
- Schwabe, L., Nader, K., & Pruessner, J. C. (2014). Reconsolidation of Human memory: Brain mechanisms and clinical Relevance. *Biological Psychiatry*, 76(4), 274–280. https://doi.org/10.1016/j.biopsych.2014.03.008
- Shippherd, J. C., & Beck, J. G. (1999). The effects of suppressing trauma-related thoughts on women with rape-related posttraumatic stress disorder. *Behaviour Research and Therapy*, 37(2), 99–112. doi:10.1016/S0005-7967(98)00136-3
- Sullivan, D. R., Marx, B., Chen, M. S., Depue, B. E., Hayes, S. M., & Hayes, J. P. (2019). Behavioral and neural correlates of memory suppression in PTSD. *Journal of Psychiatric Research*, 112, 30–37. https://doi.org/10.1016/j.jpsychires.2019.02.015
- Trenerry, M., Crosson, B., DeBoe, J., & Leber, W. (1989). *Stroop neuropsychological screening test manual*. Psychological Assessment Resources.
- Vallejo, A. G., Kroes, M. C. W., Rey, E., Acedo, M. V., Moratti, S., Fernández, G., & Strange, B. A. (2019). Propofol-induced deep sedation reduces emotional episodic memory reconsolidation in humans. *Science Advances*, 5(3), Article eaav3801. https://doi.org/10.1126/sciadv.aav3801
- van Liempt, S., van Zuiden, M., Westenberg, H., Super, A., & Vermetten, E. (2013). Impact of impaired sleep on the development of PTSD symptoms in combat veterans: A prospective longitudinal cohort study. *Depression and Anxiety*, 30(5), 469–474. https://doi.org/10.1002/da.22054
- Weathers, F. W., Litz, B. W., Herman, D. S., Huska, J. A., & Keane, T. M. (1993). *The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility*. 9th Annual Meeting of the International Society for Traumatic Stress Studies.
- Weathers, F. W., Ruscio, A. M., & Keane, T. M. (1999). Psychometric properties of nine scoring rules for the clinician-administered posttraumatic stress disorder scale. *Psychological Assessment*, 11(2), 124–133. https://doi.org/10.1037/1040-3590.11.2.124
- Wegner, D. M., Quillian, F., & Houston, C. E. (1996). Memories out of order: Thought suppression and the disturbance of sequence memory. *Journal of Personality and Social Psychology*, 71(4), 680–691. doi:10.1037/0022-3514.71.4.680
- Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology*, 51(1), 59–91. https://doi.org/10.1146/annurev.psych.51.1.59
- Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3), 225–240. https://doi.org/10.1162/105474698565686
- Zimand, E., Anderson, P. A., Rothbaum, B. O., & Hodges, L. (2001). *Immersion questionnaire*. Virtually Better, Incorporated.