Psychological Research (2024) 88:974–986
https://doi.org/10.1007/s00426-023-01909-6

RESEARCH

The effects of emotion on retrospective duration memory using virtual reality

Omran K. Safi[1] · Yiran Shi[1] · Christopher R. Madan[2] · Tyler Lin[1] · Daniela J. Palombo[1]

Received: 11 September 2023 / Accepted: 28 November 2023 / Published online: 21 December 2023

Abstract
Our memories for temporal duration may be colored by the emotions we experience during an event. While emotion generally enhances some aspects of memory, temporal duration has been shown to be particularly susceptible to emotion-induced distortions. However, prior work has faced difficulty when studying this phenomenon, having to make some trade-offs on ecological validity or experimental control. Here, we sought to bridge this gap by studying the effects of emotion on temporal duration memory using virtual reality. In the present study, a final sample of 69 participants experienced a series of negative-emotional and neutral worlds within virtual reality. Following this, participants provided ratings of emotionality (arousal, valence, pleasantness) and retrospective duration estimates (i.e., remembered time). We hypothesized that negative events would be recalled as having a greater duration than neutral events (H1). We additionally hypothesized that negative, but not neutral, events would be recalled as being longer than the true duration (H2). The results supported H1 while failing to provide evidence in support of H2. Together, the results bolster the importance of emotion, especially negative emotion, in shaping how we remember the temporal unfolding of the past.

Keywords Emotion · Episodic memory · Negative emotion · Duration memory

Introduction

Time is the thread that binds the disparate experiences of our lives into a coherent story. It is a crucial component of our episodic memories, defined as our memories for specific events from a particular place and time (Tulving, 1972). Despite such importance, our recollection of time can be distorted and dissociated from our veridical experiences. A major factor influencing our memory for time is emotion—the focus of this paper. Herein, emotion is defined as states that vary on dimensions of arousal and valence (e.g., Todd et al., 2020).[1] Within this definition, arousal refers to the degree or amount of physiological and/or subjective stimulation evoked by particular events or stimuli (i.e., calmness to excitement), while valence is conceptualized as the degree to which a particular event or stimuli is positive or negative (see review by LaBar & Cabeza, 2006 also see Todd et al., 2020). As detailed below, a broad literature has examined the influence of emotion on time and has shown that emotion distorts some aspects of time, while preserving others. The present study focuses on the role of negative emotion in temporal distortion, specifically in the context of *retrospective duration memory* (i.e., wherein time is incidentally encoded and probed in a later memory assessment).

✉ Daniela J. Palombo
daniela.palombo@psych.ubc.ca

1 Departmentt of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada

2 University of Nottingham, Nottingham, England

[1] Note that other definitions of emotion also focus on discrete states (e.g., anger, fear, happiness) or the extent to which approach/avoidance behaviors are activated (see Todd et al., 2020, for discussion).

Effects of emotion on temporal duration

Studies examining memory for temporal duration have traditionally employed the use of unfolding stimuli such as real-life events, videos or music clips (with a minority of work using words). Such work has generally found that emotional events—particularly when they are negative—are recalled as being longer in time than neutral events, although not unequivocally.[2] We discuss this work below.

In a unique study examining real-life events, Campbell and Bryant (2007) found that fear was associated with an increase in remembered time in a study of novice skydivers; excitement showed the opposite pattern. In another study, Frederickx et al. (2013) asked participants coming off an amusement park ride to provide ratings of arousal, valence, and duration. They found that only increased arousal was associated with increased duration estimates—also see results reported in Bailey and Chapman (2023). Notably, the increased arousal observed in Frederickx et al. (2013) was associated with a positively valenced event. Other studies have also examined retrospective judgments of event durations from real-life stressful events (Cardena & Spiegel, 1993; Noyes & Kletti, 1977; Ursano et al., 1999).

Some studies have been conducted using laboratory procedures. Anderson et al. (2007) found that participants estimated longer durations for clips of the 9–11 terrorist attacks compared to control clips. A similar effect was also observed by Johnson and MacKay (2019) who found that participants judged duration for taboo words to be greater than neutral words. In a study examining happy and sad musical excerpts, Bisson and colleagues (2008) found that sad musical excerpts were associated with longer remembered duration than neutral excerpts.[3] Similarly, Pollatos et al. (2014) found that fearful film clips were associated with an increase in time dilation in memory, while Ozgor et al. (2018) found this pattern both for positive and negative (versus neutral) video clips. In an older study by Loftus and colleagues (1987), participants overestimated the remembered duration of bank robbery clips and the degree of overestimation was related to the level of stress of the clip such that more stressful clips led to greater overestimation. Hence, a series of studies show that emotion tends to dilate retrospective time duration, an effect that seems to track with the arousing nature of the event in some studies—albeit more work is needed to tease apart valence and arousal effects.

Still, not all studies have found this pattern of results. Dev et al. (2022) found no difference in remembered duration estimates for a single video clip either eliciting high versus low negative emotionality, although they did observe overestimation of duration overall, consistent with some literature (also see Grondin et al., 2014; see Petrucci & Palombo, 2021 for a review). Other studies, particularly those using musical stimuli (albeit not exclusively), have also failed to show an effect of emotion on retrospective duration estimates (Kellaris & Kent, 1992; Verduyn et al., 2013; but see Bisson et al., 2008). Finally, one study wherein participants were asked to cross a bridge and then judge the duration of the walk afterwards, showed that the effect of arousal on retrospective duration memory was conflated by differences in walking time of the to-be-remembered event (Castellà et al., 2017). Collectively some, albeit not all, of the literature shows that memory of time for emotional events are distorted.

Some researchers have asked whether the observed retrospective temporal distortion (when observed) is a product of emotion's impact on perception during the experience versus a retrospective effect of emotion on memory. In a cleverly designed study, Stetson et al. (2007) asked participants for duration estimates after a free fall experience while measuring their temporal resolution (i.e., perception) during the fall. They found that participants experiencing the free fall retrospectively overestimated the duration of their fall despite there being no evidence for an increase in temporal resolution during the fall itself. This study, thus, suggests that the dilating effects of emotion on temporal duration may be due to memory and not due to a time dilation in perception (but see Johnson & MacKay, 2019).

If emotion dilates temporal duration memory, why might it do so? Some theoretical models have been proposed within the literature that might help explain this phenomenon. Ornstein (1969) introduced the idea that retrospective duration estimates are dependent upon the contents of memory—specifically, the greater the contents of memory the longer the associated duration estimate. One popular model, the contextual change model, furthers this idea by stating that elements within memory which represent contextual change are particularly important when making duration estimates (Block, 1990; Zakay & Block, 2004). Within this model, emotional information has been proposed as one such contextual change. The presence of emotional information provides additional content which may subsequently be drawn upon when retrieving an event while making a duration estimate, thus dilating the remembered duration. Another influential model, the Arousal-Biased Competition (ABC) model posits that arousal is responsible for the modulation of competing mental representations thereby enhancing memory for certain items (Mather & Sutherland, 2011). Albeit speculative, the ABC model could also provide an explanation for the role of emotion enhancing particular

[2] Some studies exist which use prospective designs (i.e., where tracking time is a task demand) and these will be considered further in the discussion.

[3] Note that this effect was observed for a relative and verbal estimation task but not a less conventional "standard" estimation task. See Bisson et al., 2008 for details of this task.

elements within memory, with arousing content providing such a contextual shift. Together, various models may lend themselves to the prediction that emotional events will be recalled as having a greater recalled duration than neutral events, due to the prioritization of emotional information providing crucial content to be drawn upon in memory when making a retrospective duration estimate.

Current study

Although there have been a number of studies to date examining the basic effects of emotion on temporal duration, the results are equivocal and more work is needed to build confidence in the prior findings and support subsequent investigation into mechanisms. Yet, the literature has faced methodological challenges. Specifically, studies to date have had to deal with a trade-off between ecologically validity and experimental control. This has made it difficult to simultaneously provide a naturally unfolding experience while attempting to control aspects of the experience. As such, most studies to date have opted to choose one side of this pendulum. While some research has focused on capturing an authentic experience, which bolsters ecological validity (Bailey & Chapman, 2023; Campbell & Bryant, 2007; Frederickx et al., 2013), those studies are more challenging to conduct and compromise some degree of experimental control. Other work has strived to retain experimental control (Özgör et al, 2018; Johnson and Mackey, 2019) but seldom reflect how time and emotion are experienced in the real world. While both lines of work have offered valuable insights into the effects of emotion on temporal duration (and the effects of emotion on memory more broadly), difficulties arise when trying to unify their findings due to the limitations of each respective paradigm.

In the present study, we sought to bridge the gap between ecological validity and experimental control in studying the effects of emotion on temporal duration memory using virtual reality (VR). Previous work has demonstrated the utility of using VR to study mnemonic phenomena (e.g., Cadet et al., 2022; McCall et al., 2022; Quezada-Scholz et al., 2022). Although VR is not akin to the real world, it can provide highly immersive naturally unfolding life-like events while retaining high experimental control (e.g., controlling for the time spent in an experience, or the stimuli presented). To achieve the goal of examining the effects of emotion on temporal duration, a novel paradigm was developed to elicit emotional responding. Participants were brought into the laboratory, where they navigated a series of negative and neutral "worlds" in VR ("micro-events"). Afterwards, in a memory test, they provided a retrospective duration estimate for each of the worlds, in addition to ratings of emotionality (as a manipulation check). In our study, negative stimuli were used due to the ease of eliciting arousal (relative to positive stimuli, which tend to produce more variable responding and thus mixed findings in the literature). We hypothesized that negative worlds will be associated with greater recalled duration than neutral worlds. Based on some prior work, we additionally hypothesized a time dilation effect such that negative, but not neutral worlds, would be recalled as having a greater duration than the true amount of time that elapsed.

Methods

Ethical statement

This study took place at the University of British Columbia with approval from the UBC Behaviour Research Ethics Board (H22-01559). All participants provided informed consent. This study was not preregistered.

Participants

Data were collected from a total of 91 undergraduate students at The University of British Columbia who were recruited through the Department of Psychology's Human Subject Pool. All participants were between the ages of 18 and 35 and had normal or corrected to normal vision. The first 12 participants were discarded from the study due to an experimental change in which we clarified the task instructions so it was clear that we were asking about the duration of the entire micro-event. At this stage, we restarted data collection (without analyzing the initial data).

Within the new sample of 79 participants, a total of 10 participants were excluded from the final sample. Three participants were excluded due to program malfunction (i.e., VR crashing or controller glitch during testing) and one participant was excluded due to not following instructions (i.e., did not comply with experimenter instructions and used their phone during the task). Two participants were excluded due to a walking time or rating time which deviated strongly from the group mean (i.e., greater than three standard deviations from the mean). Two participants were excluded for having mean duration estimates which deviated very strongly from the mean (suggesting poor task compliance). Two participants were excluded due demonstrating a systematic response pattern (i.e., putting the same duration value of "0" across all events).

The final sample consisted of 69 participants ($M_{age} = 20.43$, $SD_{age} = 2.82$, 1 genderqueer/gender non-conforming, 2 non-binary, 49 women, 17 men). Notably, there were no instances of excessive motion sickness;

consequently no exclusions were necessary on that basis. Critically, all exclusionary decisions were made prior to data analysis pertaining to testing our hypotheses. In considering power, for our first hypothesis, some prior studies that have explored the effect of emotion on duration memory have reported medium (Johnson and MacKey, 2019; Campbell & Bryant, 2007) to large effect sizes (Anderson et al., 2007). Using a medium effect size of d = 0.50, we would need 54 participants to achieve 0.95 power for a paired t-test (N = 57 for a Wilcoxon sign-ranked test) as calculated via G*Power (Faul et al., 2007). Thus, with our final sample of $N = 69$, we are well-powered to detect an effect of emotion on retrospective duration memory if one exists. We do not believe the literature affords enough information to calculate power for our second hypothesis.

Materials

Surveys

Participants completed a series of questionnaires throughout the study which were administered via Qualtrics. Specifically, participants completed a demographics and health questionnaire developed by the Memory and Imagination lab for the purpose of characterizing our sample. Participants also completed the Center for Epidemiologic Studies Depression Scale (Radloff, 1977), the Psychometric Properties of the Posttraumatic Diagnostic Scale for DSM-5 (PDS-5; Foa et al., 2016), and the State–Trait Anxiety Inventory for Adults (STAI; Spielberger, 1983). These surveys were not used in the analysis of this study and will not be discussed further.

Stimuli

The stimuli used in this study were collected online from the Unity Asset Store, Sketchfab, and the OpenVirtualObjects database (Tromp et al., 2020). A total of 80 objects were selected to be used in the study. The objects were separated into "unconditioned" (US) and "conditioned" (CS) stimuli. A US is a stimulus that evokes an emotional or neutral response without any prior learning. A CS is a stimulus that evokes emotional or neutral responding via association with the US. Such nomenclature is taken from the classical conditioning literature and is used here only for ease in describing the stimuli employed in this study and because our paradigm was used to examine ancillary hypotheses about learning (planned for a separate report but not included here). Thus, for continuity and transparency, we use this naming convention in this document but we additionally reference "Negative" (i.e., a negative US paired with a CS) and "Neutral" (i.e., a neutral US paired with a CS) Pairs for ease of readability.

There were 20 negative US (US-Negative), 20 neutral US (US-Neutral), and 40 CS, which were randomly co-paired into Negative and Neutral Pairs (to be used in the VR worlds, discussed below). See Table S1 for a complete list of the stimuli and pairings. As these stimuli did not originate from a standardized database, we conducted an independent norming study ($N = 50$; across two phases of norming; Phase 1: $N = 30$ and Phase 2: $N = 20$) prior to the beginning of the current study to ensure the validity of these stimuli. A total of 96 items were used in the second phase of the norming study (4 extra in each of above-mentioned categories) from which our final 80 were derived.[4] Participants provided ratings of arousal (0–100), valence (− 5 to 5), complexity (0–10), and pleasantness (1–10) for all the items and a rating of relatedness (1–5) for all object pairs, with higher numbers denoting a higher rating on a given scale. In addition, we conducted an image analysis to assess the low-level visual properties of our stimuli. Following these analyses, we made exclusions yielding the ranges listed in Table S2 to get to our final sample of 80 items.

Analyses reported below, which pertain to the norming data (Phase 2), include both parametric and non-parametric comparisons; the latter were used when significant normality violations were present in the data.

US-Negative ($M = -2.90$, SD = 0.87) items were rated as more negative than US-Neutral items ($M = 0.36$, SD = 0.86), ($t(38) = -11.92$, $p < 0.001$, $d = 3.77$, 95% CI [− 3.81, − 2.71]), US-Negative items ($M = 70.55$, SD = 7.93) were rated as more arousing than US-Neutral items ($M = 35.89$, SD = 10.41), ($t(38) = 11.85$, $p < 0.001$, $d = 3.75$, 95% CI [28.74, 40.58]), and US-Negative items ($M = 2.42$, SD = 0.70) were rated as less pleasant than US-Neutral items ($M = 5.49$, SD = 0.93), ($t(38) = -11.81$, $p < 0.001$, $d = 3.73$, 95% CI [− 3.59, − 2.54]). For complexity, we observed significantly higher ratings for US-Negative ($M = 6.54$, SD = 1.09) versus US-Neutral items ($M = 5.50$, $SD = 1.14$) ($t(38) = 2.93$, $p = 0.006$, $d = 0.93$, 95% CI [0.32, 1.75]); however, this complexity was not reflected in our image analysis of objective measures of complexity where we found no significant group differences (See Table S3 for details).[5] For our CS items, we found no significant differences between CS-Negative and CS-Neutral stimuli for valence, arousal, complexity, and pleasantness (See Table S4 for details), suggesting random assignment of these items to

[4] To achieve a normed set, some items were dropped and replaced in phase 1 and further norming was conducted in Phase 2 to select the final set of 80 items.

[5] Indeed, prior work suggests that emotional stimuli are rated as more complex, even when several low-level image characteristics are well matched (see Madan et al., 2018).

the two conditions was successful. We additionally observed no significant difference for the relatedness between the negative ($M = 1.36$, SD = 0.33) and neutral pairs ($M = 1.47$, SD = 0.31) ($W = 152$, $p = 0.20$).

Within our final sample of norming stimuli selected for the study, our US-Negative items had a range of 55.5–86.5 for arousal and a range of − 4.2 and − 1.5 for valence. Our US-Neutral items had a range of 17.8–53.4 for arousal and a range of − 1.4 and 2.0 for valence (see Table S2).

Our CS items fell within the range of − 0.8 to 1.4 for valence, and between 14.5 and 45.0 for arousal. These norming data are available on the Open Science Framework at: https://osf.io/qxhua/.

Virtual worlds

We created a total of 10 virtual worlds for this study. Each of the worlds was designated as negative or neutral in valence (5 per condition, repeated 4 times each; see below). "Worlds" can be conceptualized as the virtual episodes (defined as a "micro-events"; as discussed later) that participants experience (e.g., exploring a library, an office). Conditions were balanced in terms of indoor or outdoor worlds, and on the complexity of the world. The virtual worlds used in the present study were created using Unity (v. 2020.3.18f1), a video-game development engine (Unity, 2020). Participants experienced the worlds using a Vive Cosmos (HTC Corporation, Taipei, Taiwan) headset alongside the right-hand controller for movement. Refer to Table S5 for a complete list of worlds.

Procedure

The experiment was conducted in person at the Department of Psychology building on the University of British Columbia Vancouver campus. Participants began the study by providing informed consent. Participants were then provided with a brief overview of the experiment and subsequently completed the demographics and health questionnaire while the experimenter waited outside. Next, the participant was shown the virtual "play area" and trained on how to use the VR headset (see supplementary information for detailed instructions).

Before entering the virtual worlds, participants were oriented on how to use the headset and controllers. Participants were first shown how to adjust the interpupillary distance of the headset so that all visual stimuli would be clear. After this was set, participants were shown how to use the controller to move within the virtual worlds and provide ratings and make selections when prompted. Participants stood in the center of the testing room and used the joystick on the controller to walk around the virtual environments. Participants then entered the virtual world and began the tutorial portion of the experiment. The tutorial was split into two stages, a VR tutorial and an experiment tutorial (Fig. S1).

In the first phase, participants were taught how to navigate in the virtual world and make selections on slider scales when presented. The purpose of this stage was to make sure participants were comfortable navigating in virtual reality prior to beginning the experiment. At the beginning of this stage, participants also had an opportunity to double check that everything was visually clear and were allowed to re-adjust the headset if needed. Participants also received verbal instructions from the experimenter as they navigated this stage (e.g., on how to use the slider scales). In the second phase, participants received instructions for the experimental task. At the beginning, participants were informed that they will be receiving instructions for what they will be doing in the experiment and to ask questions if they are unclear on anything. At the end, they were asked to repeat the instructions back to the experimenter to ensure that they were clear on what they would be doing.

Encoding

Participants navigated micro-events (i.e., individual "episodes") where their goal was to approach a red circle with a white dome around it. Specifically, they were instructed to approach the circle where they would find two objects and they would have 10 s to examine the objects once they entered the red circle (Fig. 1; Fig. S2). Such a design ensured that all participants had the same goal on each trial (i.e., to engage the stimuli) and were not left exploring the worlds with no purpose in mind. After this, participants were moved to a rating screen and provided a rating for how easy it is to remember the object pair (Fig. S3; not analyzed).

The stimulus pairs were randomly assigned to each world within condition prior to beginning the experiment but fixed across participants. As noted, the 10 worlds (i.e., the background context) were each presented a total 4 times (once per block). The order in which the worlds were presented was randomized within each block and were unique to each participant. After completing a block, participants received a 30-s break during which they kept the headset on. Once they had completed all the blocks, the headset was taken off and the participant was given a 10-min break. Participants completed a coloring task during this break to prevent rehearsal of the stimuli encountered in the encoding portion. Participants remained standing throughout all of encoding including during the 30-s breaks between blocks but were seated during the break between encoding and retrieval (i.e., during coloring).

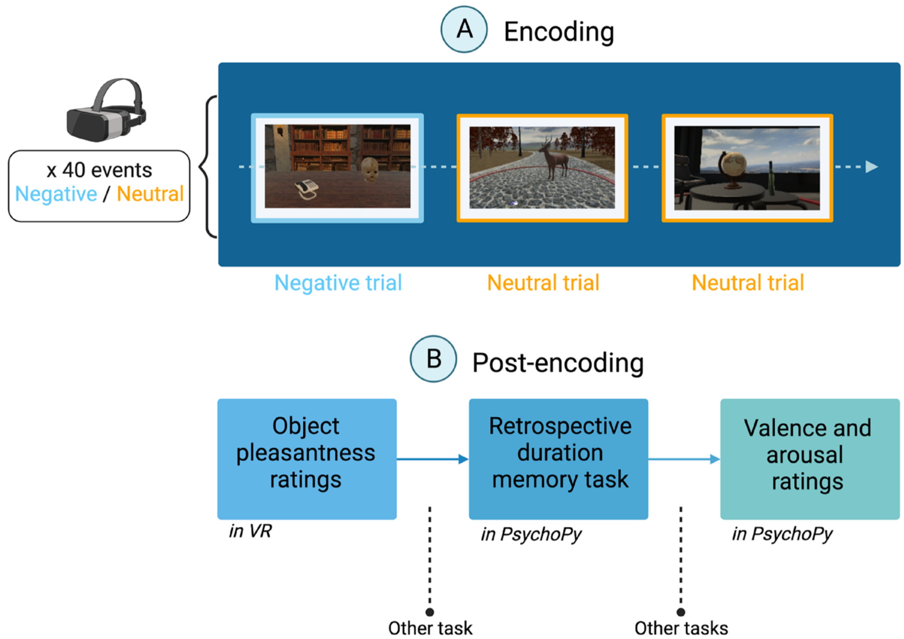

Fig. 1 Schematic of experimental paradigm. Figure created using BioRender with permission

Retrieval

After 10 min had elapsed, participants re-entered the virtual world to provide pleasantness ratings for the items encountered during encoding. The CS stimuli were presented before the US stimuli and the order of presentation was randomized within each respective CS/US category. The objects were presented in succession and participants were asked to provide a pleasantness rating on a 1 (not pleasant at all) to 10 (very pleasant) scale (Fig. S4). The purpose of this task is ancillary to the present work although the pleasantness ratings of the US stimuli were used as an additional manipulation check in the present study (see below).

Next participants took off the headset and completed a series of tasks via PsychoPy (v. 2022.1.1; Peirce et al., 2019) on an MSI GP75 Leopard laptop with a 17.3-inch display. These included a cued recall task, a retrospective duration task, a source memory task, and ratings of arousal and valence for the micro-events. The retrospective duration task and the arousal and valence ratings were the main focus of this paper and are discussed below.[6]

Retrospective duration memory

During the retrospective duration task, participants were randomly presented with an image of a micro-event and asked to provide a retrospective duration estimate. The image presented showed the object pair and the associated world to differentiate the different micro-events (Fig. S5). All the images were captured such that the objects occupied the same amount of relative space within the frame. All images were displayed in the center of the screen with each image measuring 8.5 inches by 4.5 inches. Images from the micro-events were presented in succession and participants entered a single number while responding to the question “Please rate how long this event was from the moment you entered the world to the moment you moved to the rating screen. Enter a number in seconds.” There was no time limit for responses. The retrospective duration task was presented immediately after the cued recall task so that participants would be seeing the images of the micro-events for the first time to prevent any influence of exposure on duration estimates.

Arousal and valence ratings

After completing all the aforementioned tasks, participants provided ratings of arousal and valence for each micro-event (as a manipulation check; detailed below). Just as in the retrospective duration task, participants were presented with an image for each micro-event. The image was identical to the images used in that task which displayed the object pair and the associated world. The images (measuring 8.5 inches by 4.5 inches) were randomly presented in succession at the top of the screen with the arousal and valence rating sliders below. For each micro-event, participants provided an arousal rating on a scale from 0 to 100 and a valence rating on a scale of -5 to 5. The arousal rating asked, “On a scale of

[6] At the time of this report, we have only analyzed (1) the data reported in this paper and (2) the object ratings of the CS items, which are intended for a separate manuscript; thus those data are not reported here. None of the questionnaire data have been analyzed. The cued recall and source memory data have not been analyzed.

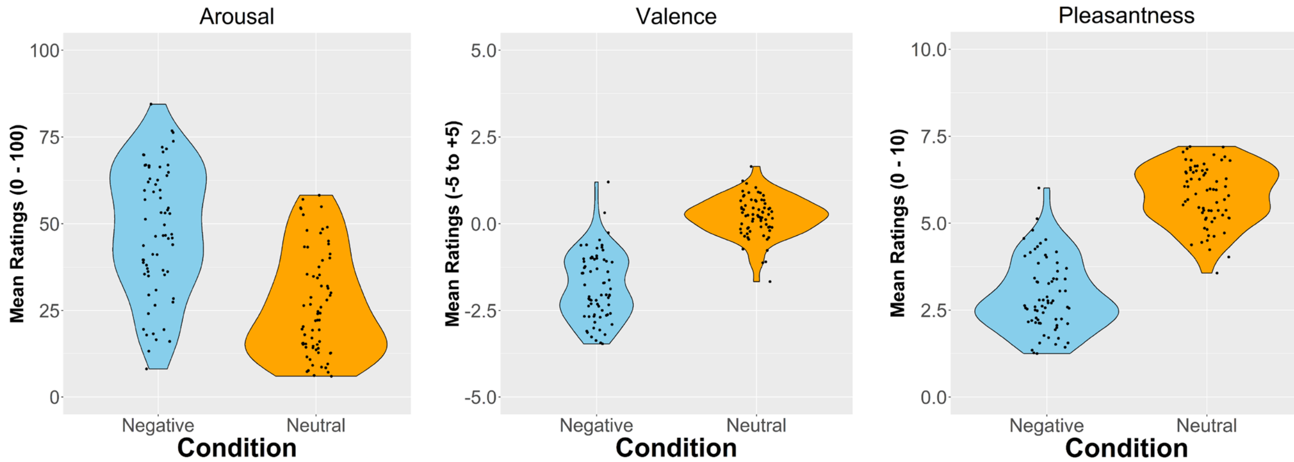

Fig. 2 Mean ratings of arousal and valence and pleasantness (the y axis displays the range of the scales) for negative and neutral conditions for each participant (black filled circles). Distributional information for each condition is shown via violin plots

0 (no arousal) to 100 (extremely arousing), how arousing did you find this event?" and the valence rating asked "Please rate how positive, negative, or neutral you found this event from -5 (very negative) to 5 (very positive)" (Fig. S6).

Results

Manipulation checks

As we were working with a novel set of stimuli and worlds, we wanted to confirm (in addition to our norming) that our intended emotional manipulations had been successful. To this end, we examined participant ratings of arousal and valence for the different micro-events. As an additional manipulation check, we examined participant ratings of pleasantness for the items they encountered while going through the worlds (as noted, the pleasantness rating was originally included for ancillary hypotheses). In what follows, paired-samples Wilcoxon tests were used when the data were non-normal; otherwise paired t-tests are reported. Note that we expected that our ratings could be somewhat diluted as a result of the repeated presentation of the stimuli as they were seen during encoding and again during the various ratings and tests (i.e., potentially dampening the emotional response to stimuli), yet as detailed below, we still observed the expected differences between conditions.

Arousal and valence ratings

For arousal, we found that participants rated negative micro-events ($M = 48.16$, $SD = 18.36$) as significantly more arousing than neutral micro-events ($M = 26.62$, $SD = 15.01$) ($V = 2352$, $p < 0.001$, $r = 0.82$) (Fig. 2). For valence, we found that participants rated negative micro-events ($M = -1.83$, $SD = 0.97$) as significantly more negative than neutral micro-events ($M = 0.2$, $SD = 0.58$) ($t(68) = -17.39$, $p < 0.001$, $d = 2.09$) (Fig. 2). Together, these results demonstrate that negative micro-events were rated as more arousing and more negative than neutral micro-events, as intended.

Pleasantness ratings

A paired-samples Wilcoxon test demonstrated a significant difference such that participants rated US-Negative items ($M = 2.89$, $SD = 1.03$) as less pleasant than US-Neutral items ($M = 5.81$, $SD = 0.85$) ($V = 1$, $p < 0.0001$, $r = 0.87$), as expected (Fig. 2).

Retrospective duration estimate

Once we had established that our negative emotion manipulation was successful, we examined participants' retrospective duration estimates for the micro-events. In addition to the participant level exclusion noted in the participants section, we examined individual trials for possible exclusion for each participant. We first examined time spent in negative versus neutral micro-events to rule out a potential confound for subsequent duration estimates. First, we z-transformed each participant's trial level time spent in the micro-events and excluded trials which were plus or minus three standard deviations from the mean. This resulted in a total of 35 trials being removed across all participants. Of these trials, 19 were negative trials and 16 were neutral trials. No single participant had more than two trials removed. We then removed the corresponding duration estimate provided for these trials

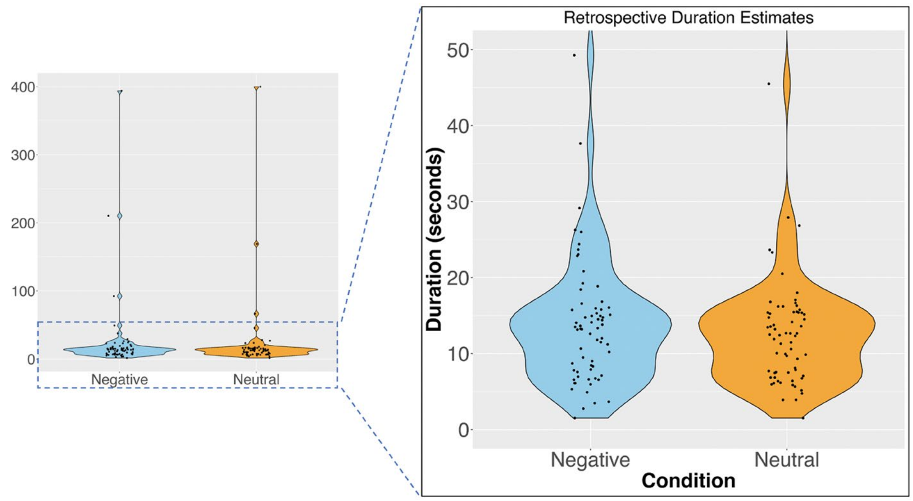

Fig. 3 Mean recalled duration estimates for negative and neutral conditions for each participant (black filled circles). Distributional information for each condition is shown via violin plots. A portion of the figure has been magnified for readability. Dotted rectangle overlay is illustrative

on the basis that an extreme "walking time" could distort duration estimates.

We found that there was no significant difference in the objective amount of time participants spent in negative ($M=31.55$, SD = 2.65) compared to neutral micro-events ($M=31.37$, SD = 2.61) ($V=1504$, $p=0.08$) (Fig. S7). We further conducted a Spearman correlation and found that the time spent in the micro-events was not predictive of subsequent duration estimates for both negative ($r=0.14$, $p=0.25$) and neutral ($r=0.18$, $p=0.14$) conditions.

Considering our critical assessment, namely, retrospective duration estimates (measured in seconds), our results show that participants recalled negative micro-events ($M=23.21$, SD = 52.52) as having elapsed for a longer amount of time than neutral ones ($M=21.15$, SD = 50.72) ($V=1698$, $p<0.001$, $r=0.42$) (Fig. 3), as predicted. Note that we were conservative with outlier removal in the retrospective duration estimate analysis. Although non-parametric statistics were used, we note that removal of two deviant participant observations notable in the data (Fig. 3; left panel) did not change the pattern of results (results are reported with those participants included). These data are available on the Open Science Framework at: https://osf.io/qxhua/.

Next, we examined the relationship between participants' duration estimates and the actual time spent in each of the micro-events.[7] For negative micro-events, we found, contrary to our second prediction, that participants' duration estimates ($M=23.21$ s, SD = 52.52) were significantly lower than the actual time ($M=31.55$ s, SD = 2.65) spent in the worlds ($V=2191$, $p<0.001$, $r=0.71$) (Fig. 4). This pattern was also found for neutral micro-events where participants' duration estimates (21.15 s, SD = 50.72) were significantly lower than actual time spent in the micro-events ($M=31.37$ s, SD = 2.61) ($V=2208$, $p<0.001$, $r=0.72$) (Fig. 4). The difference between actual time spent in the world and duration estimates was 8.34 s in the negative condition and 10.22 s in the neutral condition. Therefore, although duration estimates were extended in the negative condition, there was still an overall underestimation of time in both conditions.

Discussion

The present study employed a novel VR paradigm to explore the effects of negative emotion on retrospective temporal duration memory using naturally unfolding events. As indicated by our results pertaining to stimuli and world ratings, this paradigm successfully elicited negative-emotional responses. In line with our primary hypothesis, the results show that negative events were recalled as having elapsed for a greater amount of time than neutral ones. Importantly, this difference is not a product of differential time spent in each type of world as there was no significant difference observed between time spent in negative versus neutral worlds. However, contrary to our second hypothesis, participants recalled the duration of both negative and neutral events as shorter than the true duration of the experience.

The finding that negative-emotional events were recalled as having a longer recalled duration than neutral events is line with some literature (Loftus et al., 1987; Anderson et al., 2007; Campbell & Bryant, 2007; Bisson et al., 2008; Pollatos et al., 2014; Johnson and Makay, 2019; recently reviewed in Petrucci & Palombo, 2021). Yet, given that there have also been studies with null results (Bisson et al., 2008; Dev et al., 2022; Grondin et al., 2014; Kellaris & Kent, 1992; Verduyn et al., 2013 mixed results), our finding holds

[7] Time spent in each micro-event ("actual duration") was calculated as the time it took the participant to 'walk' to the circle (which was variable across participants) plus the 10 s they had to look at the objects (i.e., a fixed time across participants).

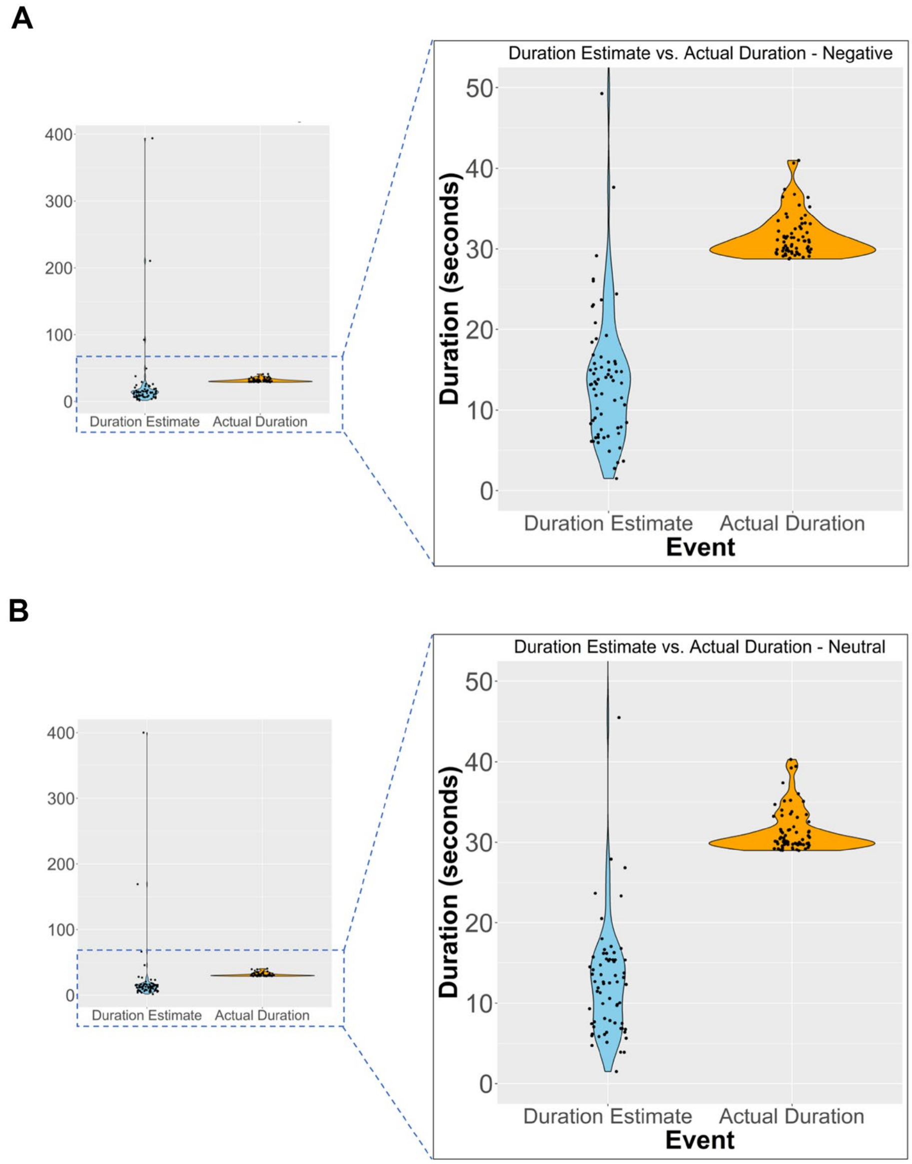

Fig. 4 Mean recalled duration estimates versus actual time spent in worlds for negative (**A**) and neutral (**B**) conditions for each participant (black filled circles). Distributional information for each condition is shown via violin plots. A portion of the figures have been magnified for readability. Dotted rectangle overlay is illustrative

significance as it provides further evidence in support of an effect of negative emotion on retrospective time estimation. It is unclear what might account for discrepancies in the literature (i.e., why some studies observe this effect and others do not), although it is notable that the pattern of results tends to be less stable in studies using music clips (but also see Dev et al., 2022). Our study, which employs a first person immersive experience, aligns with other more naturalistic or autobiographical approaches (see Introduction). Yet more work is needed to determine the boundary conditions for the effect in question.

As noted in the introduction, prior work points to the role of arousal in driving the relative temporal lengthening of remembered durations for emotional events. Given the strong ties between arousal and attention, it is conceivable that attentional mechanisms play a role in the current results. The notion that arousal and attention influence temporal judgements has been discussed both at the level of perception and memory. At the level of perception, prior work has pointed to arousal's role in guiding attentional resources toward particular features of an event which in turn affects the encoding of temporal information, particularly in prospective tasks (i.e., when participants are told in advance that they need to estimate time). Droit-Volet et al. (2020) found that perceived duration was lengthened as a function of level of arousal, such that as arousal increased so did the magnitude of the temporal distortion. In a review of this literature, Lake (2016) highlights the role of arousal and attention in different time-keeping mechanisms within pacemaker-accumulator models. Within such models, the idea is that arousal speeds up the internal time-keeping mechanism leading to an overestimation of time. Yet, some argue that attention, on the other hand, affects the collection of the information being emitted (Lake, 2016; Lake et al., 2016). The latter formulation would predict that when attention is significantly diverted away from timing information (e.g., when attention is captured by an emotionally salient distractor stimulus), an underestimation of duration could be observed (also see Lui et al., 2011). Thus, these models make nuanced predictions about how arousal and attention might together or independently affect the encoding of temporal information. Still, it is important to keep in mind that such models apply more readily to prospective tasks, where tracking time is a specific task demand. While the attentional account proposed above might predict *more* temporal compression for negative relative to neutral events, we did not observe such an effect in our retrospective paradigm. Rather, we found that the negative micro-events were recalled as being longer than the neutral ones suggesting *less* compression for negative compared to neutral events. Thus, prospective models, particularly those that focus on attention, cannot easily explain our findings.

In contrast to the often-used prospective designs, retrospective designs (such as the one used in the present study) require temporal judgements to be inferred later using the information stored in memory. Indeed, Martinelli and Droit-Volet (2022) directly compared prospective and retrospective time estimation tasks and found that the level of attention paid to time was lower in the retrospective group. As time is not an explicit demand at the moment of encoding, non-temporal information may be used when later making such judgements. Whereas prospective models sometimes tie attention and arousal to the interference of encoding temporal information, memory-based models view them as providing the information needed in retrospective judgements of time. Indeed, Ornstein (1969) argues that retrospective designs require a reconstruction thereby making duration judgments dependent on the availability of the contents of memory. Similarly, the contextual change model posits that retrospective duration judgements are made using elements available in memory, specifically elements which entail a contextual change (Block, 1990; Zakay & Block, 2004). Considering an additional popular memory theory, the Arousal-Biased Competition (ABC) theory, arousal selectively modulates competing mental representations which in turn enhances memory for certain items (Mather & Sutherland, 2011).[8] The present finding that negative micro-events were recalled as having a greater duration than neutral ones can possibly be explained by combining these theories. While temporal information was not an explicit demand at the time of encoding (i.e., when stimuli were first perceived), the effect of arousal and attention may have led to the emotionally salient events receiving a ‘boost’ in memory. When information was subsequently needed for a duration judgement, it was more readily available for the negative events compared to the neutral ones. That is, in line with some aspects of the ABC theory, the presence of increased arousal within negative micro-events may have led to greater attentional prioritization leading to subsequent enhancement of memory for emotional content. Consistent with predictions of the contextual change model, such emotional information then represents a contextual shift adding to the lengthening of the duration for the negative micro-events. Together, these models provide an avenue through which the contents of memory may have been greater for negative micro-events compared to neutral ones even when the world contained similar amounts of content.[9] Direct measures of attention and emotional enhancement of memory will be important in future work to shed light on its role in contributing to our findings and testing some of the aforementioned theories.

The second set of results were contrary to our hypothesis. Whereas we had hypothesized that negative events would be dilated, the results show that both negative and neutral events were recalled as being significantly shorter than the actual duration. Although there are instances of compression within the literature, studies examining retrospective duration have seldom observed unilateral compression across emotional and neutral conditions. Among the aforementioned studies, Johnson and MacKay (2019), Pollatos et al. (2014), and Bisson et al. (2008) all observed compression but only Bisson et al. observed compression across the board whereas the other studies only showed compression in the neutral condition and dilation in the emotional condition. One way in which our paradigm differs from previous ones is the use of VR. The use of VR may have implications for the manner in which time is remembered overall. Yet, the literature is equivocal on the role of VR in altering time estimation more broadly. Some studies, particularly those using prospective designs, have shown VR to be associated with time compression. For instance, Schneider et al. (2011) and Chirico et al. (2016) both report observing time compression when VR was used during chemotherapy. However, other work from Tobin et al. (2010) examining prospective and retrospective estimates has conversely found that both involved an overestimation of time. Moreover, van der Ham and colleagues (2019) directly tested whether the medium

8 It is important to note that both Mather and Sutherland (2011) and the broader literature emphasize a trade off in memory under some conditions, wherein only certain content is enhanced by emotion (e.g., the item eliciting emotion versus the background context). Albeit speculative, we postulate that the boost for emotional content per se may be important for driving the longer duration estimates. Such a hypothesis will be tested in later studies.

9 The negative and neutral environments were designed to be well matched in terms of the complexity of the worlds with respect to the “richness” of the environments (number of objects placed in the backgrounds, varied use of colors, etc.).

of delivery (VR versus the real world) affected duration estimates and found there to be no effect of VR. These studies raise uncertainty about the role of VR per se in driving our compression effects, but a direct comparison of our task within versus outside of VR would be needed to rule out this factor.

While the observed compression is not consistent with some prior work, the finding on its own may not be surprising from a memory perspective. Our episodic memories do not retain veridical reproductions but rather represent summary reconstructions and are therefore compressed versions of experienced events (Conway, 2009; Jeunehomme et al., 2018). Thus, it should not be surprising that the recalled duration for these events also demonstrates compression. Albeit highly speculative, our observed compression may be reflective of past work showing that our episodic memories are temporally compressed replays of the events we experience (D'Argembeau et al., 2022; Jeunehomme & D'Argembeau, 2019; Jeunehomme et al., 2018). Within this reproduction, compression is measured using experience units which are defined as recalled moments of experience. Jeunehomme et al. (2018) shows that compression is related to the density of experience units recalled. Crucially, Jeunehomme and D'Argembeau (2019) found that the density of recalled experience units was predictive of duration estimates for events in memory. Specifically, duration estimates significantly increase as the density of experience units increases. Jeunehomme et al. (2018) identify that events with more goal-directed actions and spatial displacements are accompanied with a higher density of experience units. As our paradigm involved a relatively simple task of finding and walking toward a dome it may be that a high density of units was not present when participants recalled the event (particularly given that the trial unique stimuli occurred at the latter part of the trial).

Limitations and future directions

In considering limitations of the present work, first, we only collected subjective ratings of arousal and valence as our measure of emotion and this was done retrospectively. This paradigm would benefit from the incorporation of psychophysiological measures such as skin conductance and heart rate, which can allow for a moment to moment measure of arousal throughout a micro-event. Moreover, having such data will provide an objective measure to corroborate the subjective account provided by participants. Second, although the goal of our study was not to examine mechanisms, given the importance of attention to both emotion and temporal duration the implementation of eye-tracking would be an important next step. Using eye-tracking data as proxy for attention, along with the above-mentioned emotion measures, will allow a more nuanced understanding of the relationship between these factors and could help adjudicate between the claims made by different models of emotional memory that pertain to retrospective time. Next, future work should strive to extend this paradigm by adding longer and more varied events to even better reflect the types of events we experience in everyday life. Finally, while VR provides utility as a medium for creating immersive, unfolding experiences, it also comes with its own challenges. As the effects of VR on human experience are still being explored, future work should be vigilant in its use.

Conclusion

The present study examined the effects of emotion on temporal duration memory using VR to bridge the gap between the lab and the real world and in doing so sought to extend prior literature by further investigating an emotion-induced mnemonic time dilation effect observed in the literature. Our results contribute to a larger body of work seeking to understand the effects of emotion on temporal duration. While VR cannot fully capture the nuance and complexity of the real world, representing an approximation at best, this study demonstrates (in an initial paradigm) its potential to induce emotionally salient experiences while affording a degree of flexibility not present in the assessment of real-life events or in laboratory paradigms.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00426-023-01909-6.

Acknowledgements The authors thank Emily Wang, Isabel Wilson and Noam Jacob for assistance with checking references. The authors thank Alessandra Te for assisting with analysis checking. The authors thank Theodore Yu, Akash Grewal, Jasneet Sabrao, and Christy Wong for assistance with data collection.

Author contributions Conceptualization: OKS, DJP, CRM; methodology: OKS, ES; formal analysis and investigation: OKS, TL; writing—original draft preparation: OKS; Writing—review and editing: OKS, DJP, CRM; funding acquisition: DJP; resources: DJP; supervision: DJP, CR.

Funding This work was supported by an NSERC Discovery Grant (RGPIN-2019–04596) and the John R. Evans Leaders Fund from the Canadian Foundation for Innovation (38817).

Availability of data and materials Data are available on the Open Science Framework at: https://osf.io/qxhua/. The detailed instructions provided to participants are provided in the supplementary information (SI). Additional materials are available upon request.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval This study took place at the University of British Columbia with approval from the UBC Behaviour Research Ethics Board (H22-01559). All participants provided informed consent.

References

Anderson, M. J., Reis-Costa, K., & Misanin, J. R. (2007). Effects of September 11th terrorism stress on estimated duration. *Perceptual and Motor Skills, 104*(3), 799–802.

Bailey, K., & Chapman, P. (2023). The best of times, the worst of times: Memory for emotional experiences. In A. D. Smith (Ed.), *Cognition in the Real World* (pp. 208–245). Oxford University Press.

Bisson, N., Tobin, S., & Grondin, S. (2008). Remembering the duration of joyful and sad musical excerpts: Assessment with three estimation methods. *NeuroQuantology*. https://doi.org/10.14704/nq.2009.7.1.206

Block, R. A. (1990). Models of psychological time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 1–35). Erlbaum.

Cadet, L. B., Reynaud, E., & Chainay, H. (2022). Memory for a virtual reality experience in children and adults according to image quality, emotion, and sense of presence. *Virtual Reality, 26*(1), 55–75.

Campbell, L. A., & Bryant, R. A. (2007). How time flies: A study of novice skydivers. *Behaviour Research and Therapy, 45*(6), 1389–1392.

Cardena, E., & Spiegel, D. (1993). Dissociative reactions to the San Francisco Bay Area earthquake of 1989. *American Journal of Psychiatry, 150*(3), 474–478.

Castellà, J., Cuello, C., & Sanz, A. (2017). Does time fly 20 m above the ground? Exploring the role of affective response on time perception in a high-risk sport. *Applied Cognitive Psychology, 31*(6), 644–652.

Chirico, A., D'Aiuto, M., Pinto, M., Milanese, C., Napoli, A., Avino, F., & Lucidi, F. (2016). *The elapsed time during a virtual reality treatment for stressful procedures. A pool analysis on breast cancer patients during chemotherapy. Intelligent Interactive Multimedia Systems and Services 2016* (pp. 731–738). New York: Springer International Publishing.

Conway, M. A. (2009). Episodic memories. *Neuropsychologia, 47*(11), 2305–2313.

D'Argembeau, A., Jeunehomme, O., & Stawarczyk, D. (2022). Slices of the past: How events are temporally compressed in episodic memory. *Memory, 30*(1), 43–48.

Dev, D. K., Wardell, V., Checknita, K. J., Te, A. A., Petrucci, A. S., Le, M. L., & Palombo, D. J. (2022). Negative emotion enhances memory for the sequential unfolding of a naturalistic experience. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1037/mac0000015

Droit-Volet, S., El-Azhari, A., Haddar, S., Drago, R., & Gil, S. (2020). Similar time distortions under the effect of emotion for durations of several minutes and a few seconds. *Acta Psychologica, 210*, 103170.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Foa, E. B., McLean, C. P., Zang, Y., Zhong, J., Powers, M. B., Kauffman, B. Y., & Knowles, K. (2016). Psychometric properties of the Posttraumatic Diagnostic Scale for DSM–5 (PDS–5). *Psychological Assessment, 28*(10), 1166.

Frederickx, S., Verduyn, P., Koval, P., Brans, K., Brunner, B., Laet, I. D., & Hofmans, J. (2013). The relationship between arousal and the remembered duration of positive events. *Applied Cognitive Psychology, 27*(4), 493–496.

Grondin, S., Laflamme, V., Bisson, N., & Désautels, F. (2014). The delay before recall changes the remembered duration of 15-minute video sequences. *Applied Cognitive Psychology, 28*(5), 677–684.

Haas, J. K. (2014). A history of the unity game engine.

Jeunehomme, O., & D'Argembeau, A. (2019). The time to remember: Temporal compression and duration judgements in memory for real-life events. *Quarterly Journal of Experimental Psychology, 72*(4), 930–942.

Jeunehomme, O., Folville, A., Stawarczyk, D., Van der Linden, M., & D'Argembeau, A. (2018). Temporal compression in episodic memory for real-life events. *Memory, 26*(6), 759–770.

Johnson, L. W., & MacKay, D. G. (2019). Relations between emotion, memory encoding, and time perception. *Cognition and Emotion, 33*(2), 185–196.

Kellaris, J. J., & Kent, R. J. (1992). The influence of music on consumers' temporal perceptions: Does time fly when you're having fun? *Journal of Consumer Psychology, 1*(4), 365–376.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*(1), 54–64.

Lake, J. I. (2016). Recent advances in understanding emotion-driven temporal distortions. *Current Opinion in Behavioral Sciences, 8*, 214–219.

Lake, J. I., LaBar, K. S., & Meck, W. H. (2016). Emotional modulation of interval timing and time perception. *Neuroscience and Biobehavioral Reviews, 64*, 403–420.

Loftus, E. F., Schooler, J. W., Boone, S. M., & Kline, D. (1987). Time went by so slowly: Overestimation of event duration by males and females. *Applied Cognitive Psychology, 1*(1), 3–13.

Lui, M. A., Penney, T. B., & Schirmer, A. (2011). Emotion effects on timing: Attention versus pacemaker accounts. *PLoS ONE, 6*(7), e21829.

Madan, C. R., Bayer, J., Gamer, M., Lonsdorf, T. B., & Sommer, T. (2018). Visual complexity and affect: Ratings reflect more than meets the eye. *Frontiers in Psychology, 8*, 2368.

Martinelli, N., & Droit-Volet, S. (2022). Judgment of duration and passage of time in prospective and retrospective conditions and its predictors for short and long durations. *Scientific Reports, 12*(1), 22241.

Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science, 6*(2), 114–133.

McCall, C., Schofield, G., Halgarth, D., Blyth, G., Laycock, A., & Palombo, D. J. (2022). The underwood project: A virtual environment for eliciting ambiguous threat. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-02002-3

Noyes, R., & Kletti, R. (1977). Depersonalization in response to life-threatening danger. *Comprehensive Psychiatry, 18*, 375–384.

Ornstein, R. E. (1969). *On the experience of time*. Penguin.

Özgör, C., Şenyer Özgör, S., Duru, A. D., & Işoğlu-Alkaç, Ü. (2018). How visual stimulus effects the time perception? The evidence from time perception of emotional videos. *Cognitive Neurodynamics, 12*(4), 357–363.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Petrucci, A. S., & Palombo, D. J. (2021). A matter of time: How does emotion influence temporal aspects of remembering? *Cognition and Emotion, 35*(8), 1499–1515.

Pollatos, O., Laubrock, J., & Wittmann, M. (2014). Interoceptive focus shapes the experience of time. *PLoS ONE, 9*(1), e86934.

Quezada-Scholz, V. E., Laborda, M. A., San Martín, C., Miguez, G., Alfaro, F., Mallea, J., & Díaz, F. (2022). Cued fear conditioning

in humans using immersive Virtual Reality. *Learning and Motivation, 78*, 101803.
Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385–401.
Schneider, S. M., Kisby, C. K., & Flint, E. P. (2011). Effect of virtual reality on time perception in patients receiving chemotherapy. *Supportive Care in Cancer, 19*, 555–564.
Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the State-Trait Anxiety Inventory. Consulting Psychologists Press.
Stetson, C., Fiesta, M. P., & Eagleman, D. M. (2007). Does time really slow down during a frightening event? *PLoS ONE, 2*(12), e1295.
Tobin, S., Bisson, N., & Grondin, S. (2010). An ecological approach to prospective and retrospective timing of long durations: A study involving gamers. *PLoS ONE, 5*(2), e9271.
Todd, R. M., Miskovic, V., Chikazoe, J., & Anderson, A. K. (2020). Emotional objectivity: Neural representations of emotions and their interaction with cognition. *Annual Review of Psychology, 71*, 25–48.
Tromp, J., Klotzsche, F., Krohn, S., Akbal, M., Pohl, L., Quinque, E. M., & Gaebler, M. (2020). OpenVirtualObjects: An open set of standardized and validated 3D household objects for virtual reality-based research, assessment, and therapy. *Frontiers in Virtual Reality, 1*, 611091.
Tulving, E. (1972). Episodic and semantic memory. *Organization of memory* (pp. 381–403). Academic Press.
Unity (v. 2020.3.18f1). (2020). [Computer software]. Unity Technologies.
Ursano, R., Fullerton, C., Epstein, R., Crowley, B., Vance, K., Kao, T.-C., et al. (1999). Peritraumatic dissociation and posttraumatic stress disorder following motor vehicle accidents. *American Journal of Psychiatry, 156*, 1808–1810.
van der Ham, I. J., Klaassen, F., van Schie, K., & Cuperus, A. (2019). Elapsed time estimates in virtual reality and the physical world: The role of arousal and emotional valence. *Computers in Human Behavior, 94*, 77–81.
Verduyn, P., Tuerlinckx, F., & Van Gorp, K. (2013). Measuring the duration of emotional experience: The influence of actual duration and response format. *Quality and Quantity, 47*, 2557–2567.
Zakay, D., & Block, R. A. (2004). Prospective and retrospective duration judgments: An executive-control perspective. *Acta Neurobiologiae Experimentalis, 64*(3), 319–328.