Working memory capacity for continuous events:

The root of temporal compression in episodic memory?

Nathan Leroy [a], Steve Majerus [a], Arnaud D'Argembeau [a]

[a] University of Liège, Belgium

Accepted Manuscript (April 1, 2024)

The published article:

Authors Note

Nathan Leroy, https://orcid.org/0000-0003-2211-1345, Psychology and Neuroscience of Cognition Research Unit, University of Liège, Belgium; Steve Majerus, https://orcid.org/0000-0002-8206-8097, Psychology and Neuroscience of Cognition Research Unit, University of Liège, Belgium; Arnaud D'Argembeau, https://orcid.org/0000-0003-3618-9768, Psychology and Neuroscience of Cognition Research Unit, University of Liège, Belgium.

Correspondence concerning this article should be addressed to Nathan Leroy, Psychology and Neuroscience of Cognition Research Unit, University of Liège, Place des Orateurs 1 (B33), 4000 Liège, Belgium. E-mail: nathan.leroy@uliege.be

Author contributions

Nathan Leroy: Conceptualization, Methodology, Data curation, Formal analysis, Writing - original draft.

Steve Majerus: Methodology, Supervision, Writing - review & editing

Arnaud D'Argembeau: Conceptualization, Methodology, Supervision, Writing - original draft

Declaration of interest

Declarations of interest: none

Transparency and openness

The design of this study, the analysis plan, and all hypotheses were preregistered on OSF

(https://osf.io/4hxzs). We report how we determined our sample size, all data exclusions,

transformations, and all measures. All data, analysis scripts, and research materials are available

at https://osf.io/zkwf2.

Abstract

Remembering the unfolding of past episodes usually takes less time than their actual duration. In this study, we evaluated whether such temporal compression emerges when continuous events are too long to be fully held in working memory. To do so, we asked 90 young adults to watch and mentally replay video clips showing people performing a continuous action (e.g., turning a car jack) that lasted 3, 6, 9, 12, or 15 s. For each clip, participants had to carefully watch the event and then to mentally replay it as accurately and precisely as possible. Results showed that mental replay durations increased with event duration but in a non-linear manner: they were close to the actual event duration for short videos (3-9 s), but significantly smaller for longer videos (12 and 15 s). These results suggest that working memory is temporally limited in its capacity to represent continuous events, which could in part explain why the unfolding of events is temporally compressed in episodic memory.

Keywords: Temporal compression, working memory, event duration, episodic memory

**Working memory capacity for continuous events: the root of temporal compression in episodic memory?**

1. Introduction

Humans are endowed with the fascinating ability to mentally travel in time. This faculty allows us, among other things, to mentally relive episodes from our personal past (Conway, 2009; D'Argembeau, 2020; Tulving, 2002). Nevertheless, our memories are not like continuous video recordings. Events are mentally replayed as a sequence of experience units—each unit representing a moment of the past experience—that are separated by temporal discontinuities (i.e., moments that are not remembered; D'Argembeau et al., 2022). Because of these discontinuities, the time taken to mentally re-experience an event is typically shorter than its actual duration, a phenomenon referred to as the temporal compression of events in episodic memory (Jeunehomme & D'Argembeau, 2019, 2023). However, the cognitive mechanisms underlying this compression, particularly the factors leading to temporal discontinuities in event representation, are not fully understood. Here, we aim to shed light on the mechanisms of construction of the experience units that constitute episodic memories by investigating how continuous events of various length are maintained in working memory (WM).

A consistent body of work suggests that the experience units that compose episodic memories are formed through cognitive processing taking place during event perception. In daily life, we constantly segment the continuous flow of experience into meaningful units (i.e., events and sub-events) that are temporally delimited by the perception of event boundaries (i.e., the end of an event and the beginning of another; Kurby & Zacks, 2008; Radvansky & Zacks, 2017). Between event boundaries, WM enables us to construct and maintain a mental model of the current situation (i.e., an event model) while continuously processing incoming information

(Loschky et al., 2020; Radvansky, 2017; Richmond & Zacks, 2017; Zacks, 2020). When the situation changes (i.e., when an event boundary is perceived), the current event model is updated and its previous version is integrated in the long-term memory representation of the ongoing sequence of events (Bailey et al., 2017; Bird, 2020; Loschky et al., 2020; Lu et al., 2022).

Based on this work on event segmentation, we suggest that temporal discontinuities within episodic memories may be a byproduct of the role of WM in constructing events models. When new information needs to be continuously processed, WM capacity is temporally limited: without refreshing, the information just perceived is only maintained for a limited amount of time (Barrouillet et al., 2004). So, in the context of event model formation (during which incoming information must be continuously accumulated), the temporal limit of WM (i.e., the time during which information can be held in WM without active refreshing) should correspond to the maximal duration of an event model that can be fully maintained in mind. When this temporal capacity is exceeded before the perception of an event boundary, the experience unit formed at the end of the current event would only partially represent its unfolding, leading to temporal discontinuities in the long-term memory representation into which this experience unit is integrated (Figure 1).

Figure 1

Hypothesized mechanism behind temporal discontinuities in episodic memories
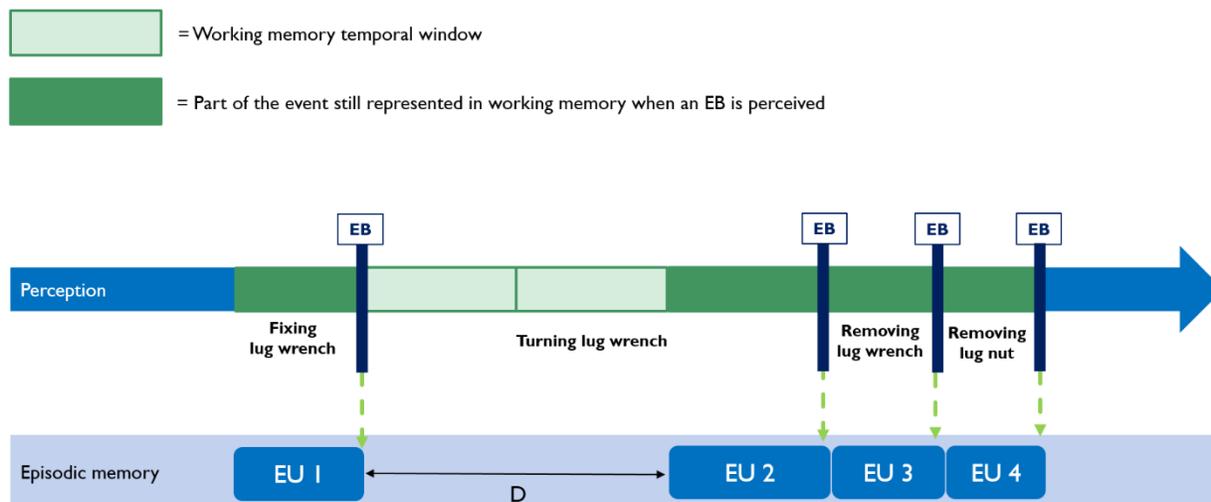
Note. The cognitive system continuously constructs memory representations of everyday activities (e.g., changing a car tire) by dividing them into discrete meaningful units (i.e., events; e.g., fixing the lug wrench, turning it, removing it, and so on). When the duration of some events exceeds WM temporal limit, their unfolding is incompletely represented, leading to temporal discontinuities in the episodic memories in which they are integrated. EB: Event Boundary; EU: Experience Unit; D: Discontinuity.

Research on temporal cognition suggests that the "subjective present" lasts about 3 s, which would correspond to the elementary units of the flow of consciousness (Fairhall et al., 2014; Monfort et al., 2020; Montemayor & Wittmann, 2014; Pöppel, 1997). WM would then enable us to maintain several of these units of "now" in an active state to form more complex event representations (i.e., event models; Richmond & Zacks, 2017; Wittmann, 2016). According to some authors, this maintenance mechanism would have an upper temporal limit situated around 10-12 s (Jeneson & Squire, 2012; Wittmann, 2016). As a result, temporal compression in

memory could emerge when individuals are faced with continuous events lasting longer than 10-12 s.

To test this hypothesis and determine the upper limit of WM capacity for continuous events, we measured the time taken by participants to mentally replay short videos depicting a single event (without event boundaries) lasting 3, 6, 9, 12 or 15 s. We expected remembering duration (the time taken by participants to mentally replay a just-seen video) to closely follow event duration for 3- to 9-s events, but then to increase less steeply for longer events (i.e., 9- to 15-s events). In other words, we predicted that the ratio between the time needed to mentally replay a video and its original duration would decrease with video length in a non-linear way: the duration of mental replay should be close to the actual stimulus duration for short videos (3 to 9 s), but smaller for longer videos (12 or 15 s). This pattern of results would suggest that the temporal limit of WM for continuous events was attained between 9 and 12 s. Beyond that duration, the maintained event model would no longer represent the entire unfolding of the current event and thus temporal compression would emerge. These hypotheses, the study design, and analysis plan were preregistered in OSF at https://osf.io/4hxzs.

2. Method

2.1. Participants

Ninety participants (49 females, 39 males, and 2 undefined) aged between 18 and 35 years ($M = 22.3$, $SD = 3.05$) were recruited through announcements on social media and word-of-mouth. The targeted sample size for this study was determined a priori with a power analysis based on Monte-Carlo simulations (Brysbaert & Stevens, 2018; DeBruine & Barr, 2021). This analysis indicated that a sample size of 90 participants would provide a statistical power of at least 90% for each of the planned analyses (see Supplementary Materials for more details). To be eligible, participants had not to be currently taking any medication that could affect their ability to concentrate or have a history of psychiatric, psychological, or neurological disorders. All participants were fluent French speakers. They all provided written informed consent, and the study was approved by the ethics committee of the Faculty of Psychology of the University of Liège (ref. 2122-009).

2.2. Materials

The stimuli were constructed based on 25 videos each showing one (or several) person(s) performing a continuous action (e.g., turning a car jack) for at least 15 s (without interruption). Some of the videos were selected from databases linked to previous studies on event segmentation (Eisenberg & Zacks, 2016; Kurby & Zacks, 2011; Sargent et al., 2013; Smith et al., 2020, 2021; Wahlheim et al., 2022) and other videos were downloaded from a website hosting free-to-use audiovisual content (https://pixabay.com/). Then, each video was edited to construct versions of the event that lasted 3, 6, 9, 12, and 15 s. The 3-s version corresponded to the first 3 s of the event, the 6-s version correspond to the first 6 s of the event, and so on. In total, we

constituted five sets of stimuli (each containing 25 videos: 5 videos for each of the 5 durations), such that the assignment of an event to a given duration was counterbalanced across participants. Each participant was randomly assigned to one of these sets (with replacement). In each set, the 25 videos were divided into 5 blocks, each block including one video of each duration. The order of blocks, as well as trials within blocks, were fully randomized.

2.3. Procedure

Testing took place in a soundproof room. Participants performed the task on a laptop via the Gorilla platform (https://gorilla.sc/). The experiment lasted 60 min on average.

Temporal compression task. In total, 25 trials (5 blocks x 5 trials per block) were presented to each participant. Each trial started with a fixation cross of 2 s, followed by a video. Participants had to carefully watch and then to mentally replay the video, as accurately and precisely as possible (as if they were watching the video again in their minds; Figure 2). The time taken by participants to mentally replay the unfolding of the video was measured by asking them to press the spacebar to indicate the beginning and end of their mental replay (Jeunehomme & D'Argembeau, 2019). To familiarize participants with the procedure, they performed 5 practice trials (one for each duration) before starting the main task (using videos that were not used in the main task). The detailed instructions are available in OSF at https://osf.io/zkwf2.

Figure 2

Unfolding of a trial in the mental replay task
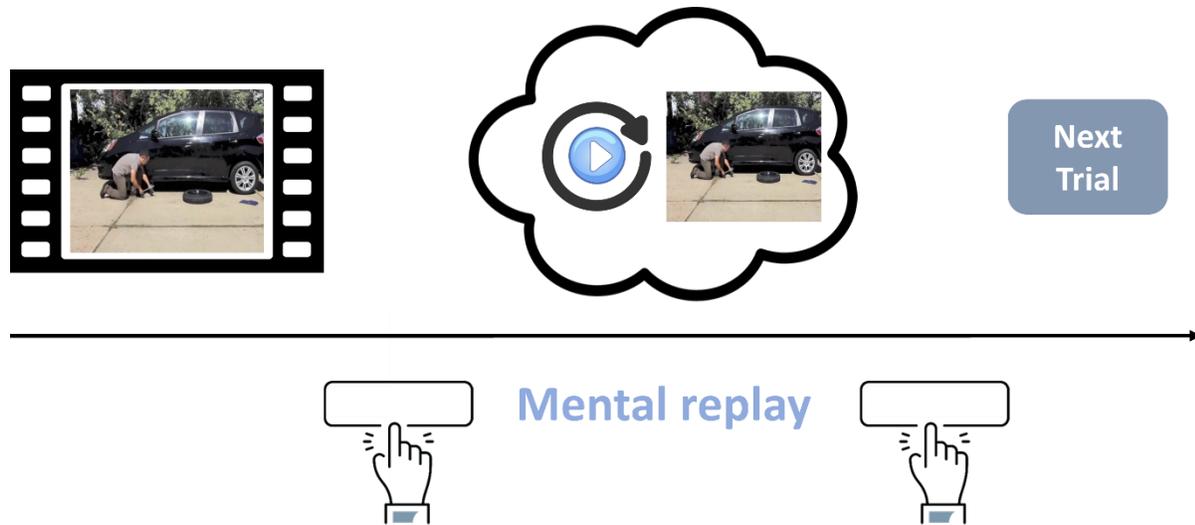
Note. Each trial started with a video clip showing a person engaged in a continuous action (performed without interruption) during 3, 6, 9, 12 or 15-s. As soon as the video ended, participants had to mentally replay its unfolding in as much detail as possible (as if they saw again the depicted event). They were asked to press the spacebar when they started and when they finished their mental replay.

2.4. Statistical analyses

Based on our pre-registered inclusion and data cleaning criteria, 17 participants were excluded and replaced by other participants (see Supplementary Materials for more details). The analyses reported here are based on 2212 observations from 90 participants[1]. Following our pre-registered analyses plan, we tested our hypotheses using linear mixed-effects models (growth curve analyses; Curran et al., 2010; Mirman, 2014). For each participant, remembering duration (RD) and temporal ratio (the ratio between RD and the actual video duration) were averaged by duration. First, we examined the relation between stimulus duration and the time taken by participants to mentally replay the videos by fitting a model with RD as outcome and a first and a second order polynomial transformation of stimulus duration as predictors. Next, we examined the relation between stimulus duration and temporal ratio in a second model with the same predictors and temporal ratio as outcome. As the conditions of application of classical linear mixed-effects models were not fully met (see https://osf.io/zkwf2 for the complete assessment), we used a robust alternative (i.e., DAStau estimator; Koller, 2013, 2016; Mason, 2022)[2]. For the two models, the random effect structure was determined following the likelihood ratio test backward selection heuristic (Matuschek et al., 2017). The model predicting RD included two random effects: a random intercept for participants, and a random slope for the linear term. The

---

[1] To ensure that the observed results were not dependent on our inclusion and data cleaning criteria, we ran again all the analyses with all data (N participants = 107, N observations = 2675). Conclusions regarding our hypotheses remained unchanged (results are available in the Supplementary Materials).

[2] The results obtained with the classical and robust analyses were quite similar and conclusions regarding our hypotheses were the same. We therefore only report here the results obtained with the robust estimates, while the results of the classical analyses are available in the Supplementary Materials.

model predicting temporal ratio included three random effects: a random intercept for participants, and random slopes for the linear and quadratic terms. We assessed the statistical significance of parameter estimates with confidence intervals (*CI*s) and *p*-values (considering an alpha of 0.05, two tailed) obtained from standard errors (*SE*) and *t*-statistics of the model using the normal approximation (i.e., treating the *t*-value as a *z*-value; Baayen et al., 2008; Mason, 2022; Mirman, 2014). Estimated RD and temporal ratio, associated pointwise standard errors and Wald's 95% CIs were computed for each event duration using variance-covariance matrix of the model.

All the analyses were performed using R (version 4.2.2; R Core Team, 2022) and RStudio (version 3.0.386; Posit team, 2023). See Supplementary Materials for more details about R packages that were used.

3. Results

First, we examined the evolution of mental replay durations (RD) as a function of event duration. Figure 3 shows the averaged RD of each participant for each duration, as well as the model estimates and their 95% CIs (detailed descriptive statistics are available in the Supplementary Materials). As expected, a growth curve analysis revealed a non-linear increase of RD as a function of event duration (linear term: $b = 6.89$, $SE = 0.26$, $95\% CI$ [6.39, 7.40], $t = 26.87$, $p < .001$; quadratic term: $b = -0.74$, $SE = 0.08$, $95\% CI$ [-0.89, -0.59], $t = -9.71$, $p < .001$). Indeed, the increase in RD was lower as event duration increased (e.g., the increase in RD between 3- and 6-s events was greater than the increase in RD between 12- and 15-s events; see Table 1, Figure 3). Importantly, RD was close to the actual event duration for short events (i.e., 3- and 6-s events) but significantly shorter than the actual event duration for long events (12- and 15-s events; see Table 1, Figure 3).

Figure 3

Increase of mental replay duration as a function of event duration
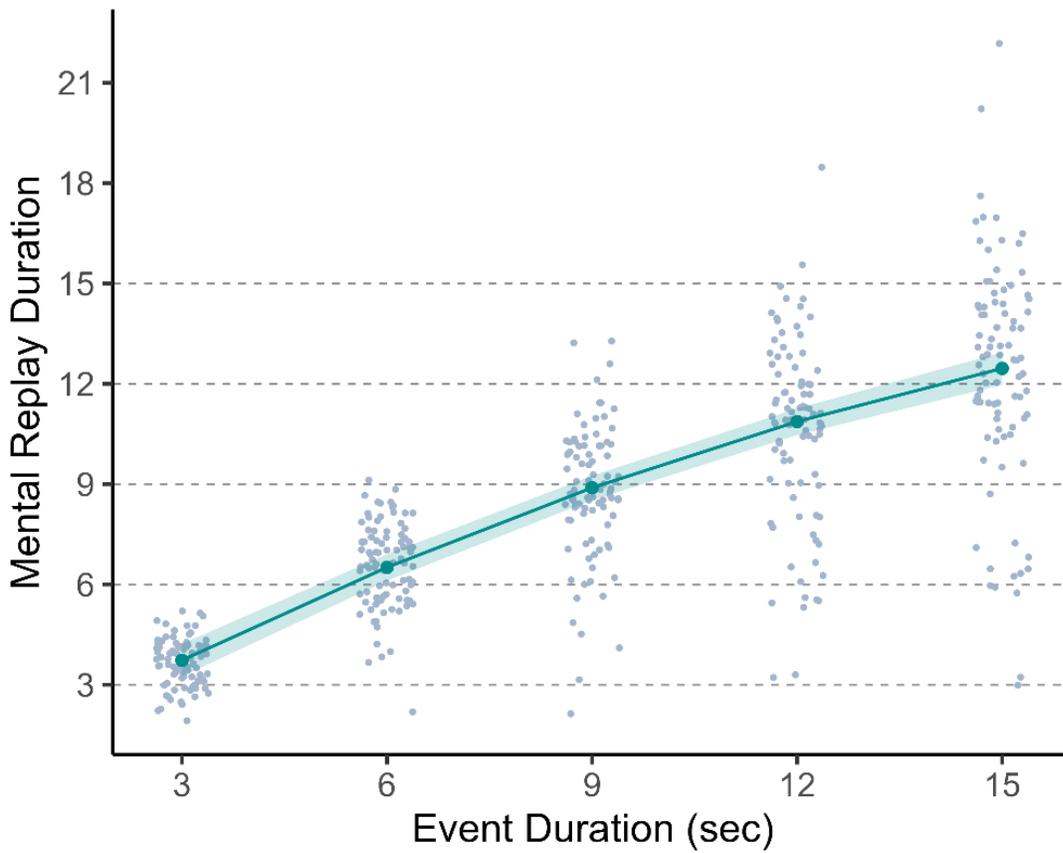
Note. The gray dots represent the observed values. Each point corresponds to the average RD of a given participant for a given duration. The green dots and the ribbon surrounding them represent the model estimates and their 95% CIs for each event duration.

To further investigate this non-linear relation between event and mental replay durations, we analyzed temporal ratios (i.e., the ratio between the time needed to mentally replay a video and its original duration; see Figure 4). As predicted, a growth curve analysis revealed a non-linear decrease of temporal ratio as a function of event duration (linear term: b = -0.32, SE = 0.02, 95%CI [-0.36, -0.27], t = -14.34, p < .001; quadratic term: b = 0.05, SE = 0.01, 95%CI [0.02, 0.07], t = 3.79, p < .001)[3].

Table 1

Growth Curve Analyses (robust linear mixed-effects models): estimates across event durations

Event Duration	Remembering Duration		Temporal Ratio	
	Estimate	95%CI	Estimate	95%CI
3 s	3.73	[3.25, 4.22]	1.23	[1.18, 1.28]
6 s	6.51	[6.11, 6.90]	1.09	[1.05, 1.14]
9 s	8.89	[8.52, 9.25]	0.98	[0.94, 1.02]
12 s	10.87	[10.47, 11.26]	0.89	[0.85, 0.94]
15 s	12.46	[11.97, 12.94]	0.83	[0.78, 0.88]

Note. For each event duration: estimated RD (left) and temporal ratio (right), with associated 95% confidence intervals.

---

[3] Model comparisons performed on non-robust growth curves (see Supplementary Materials) revealed that a model including both linear and quadratic terms provided a significantly better fit to the data than a simpler model that only included the linear term. This was the case for both the model predicting RD ($\chi^2(1)$ = 75.69, p < 0.01) and the model predicting the temporal ratio ($\chi^2(1)$ = 12.33, p < 0.01).

Importantly for our purpose, model estimates showed that mental replay durations became significantly shorter than the actual event durations when events lasted more than 9 s (i.e., temporal ratios were significantly lower than 1 only for 12- and 15-s events; see Table 1). These results suggest that continuous events start to become temporally compressed in WM representations when their duration is about 10-12 s (see also Figures 3 and 4)[4].

It is also worth noting that the estimated RD for short events (i.e., 3- and 6-s events) was slightly longer than the actual event duration (temporal ratio > 1). As a result, one could argue that the variations in temporal ratios we observed as a function of event duration could be explained by a central tendency bias. However, more detailed analyses of the evolution of RD across trials do not support this interpretation (see Supplementary Materials).

---

[4] Although temporal ratios decreased with event duration in almost all participants, the magnitude of this decrease, as well as the event duration at which it started, varied substantially between individuals (see Supplementary Materials). If, as hypothesized, the emergence of temporal compression is a byproduct of WM temporal limit, variations in WM forgetting rate should explain these individual differences. Future studies could test this hypothesis by independently measuring individual differences in WM capacity and by examining the associations between WM capacity and the event-duration-dependent decrease of temporal ratios.

Figure 4

Decrease of temporal ratio as a function of event duration
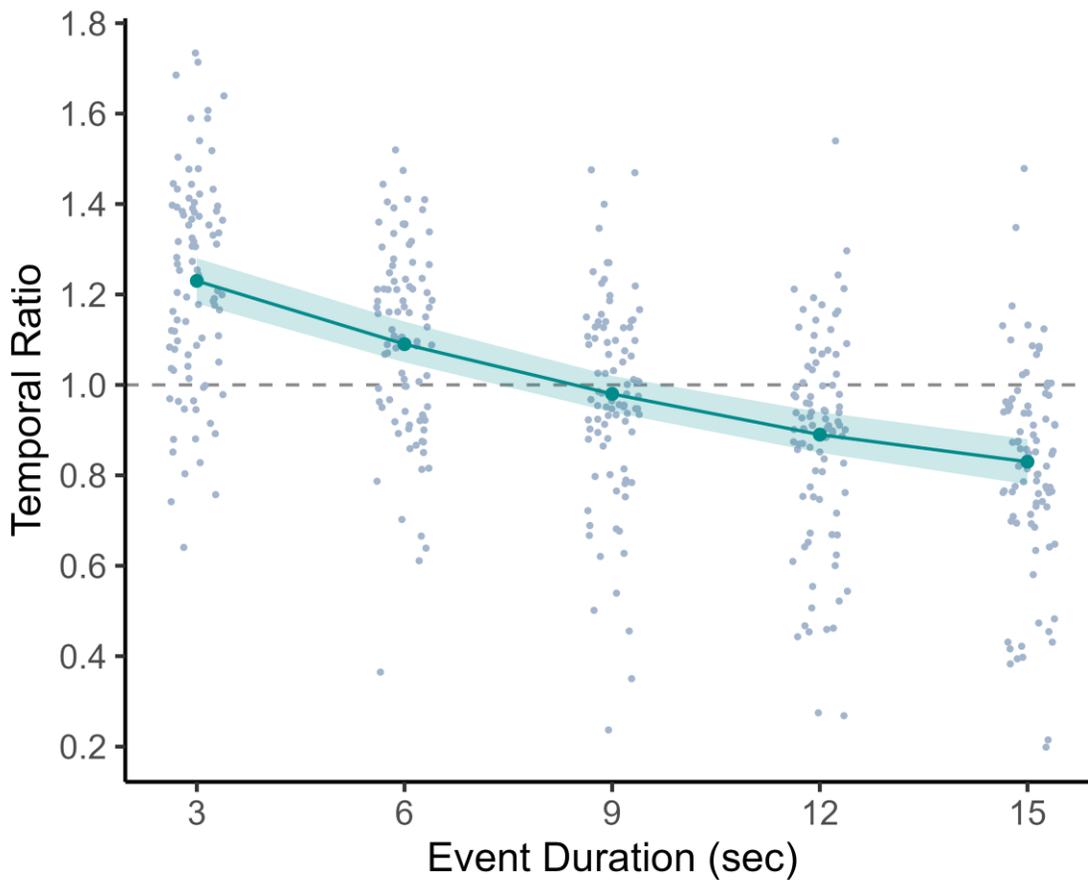
Note. The gray dots represent the observed values. Each point corresponds to the average temporal ratio of a given participant for a given duration. The green dots and the ribbon surrounding them represent the model estimates and their 95% CIs for each event duration.

4. Discussion

Episodic memories represent past events as sequences of experience units (i.e., segments of prior experience) that include discontinuities, such that the unfolding of events is temporally compressed (D'Argembeau et al., 2022). Previous studies suggest that the compression rate of an event depends on the number and duration of experience units representing its unfolding in memory (Jeunehomme & D'Argembeau, 2020, 2023). In the current study, we aimed to shed light on the mechanisms of construction of experience units by determining at what point the unfolding of continuous events that do not include event boundaries starts to be temporally compressed in memory. Using a novel paradigm that assesses WM for continuous events, we found that, as expected, the temporal compression of continuous events held in WM is not systematic but emerges when these events exceed a given duration: RD was significantly shorter than the actual event duration only for 12- and 15-s videos. These results suggest the existence of a temporal limit of WM in representing continuous events (i.e., the maximum duration up to which the stream of current experience can be accumulated and integrated into a single event model), which may in turn determine the duration of experience units that are integrated in episodic memory.

We intentionally used continuous events that do not contain event boundaries to investigate at what point the stream of experience starts to be temporally compressed in the current event model maintained in WM. Our results suggest that the information accumulation process underlying event model formation (see Güler et al., 2023; Lu et al., 2022) can be "lossy" because of WM capacity limits in representing continuous events. This, in turn, would lead to the formation of incomplete episodic memory representations when the time elapsed between two event boundaries is too long (Jeunehomme & D'Argembeau, 2020, 2023). Nevertheless, factors other than event duration per se could also determine when an event model begins to be temporally compressed. When we are faced with objects or agents that move and change, our perceptual system divides the continuous sensory information into discrete perceptual units (e.g., a person raising the left arm; Wood, 2007).[5] Like verbal information chunks, the number of perceptual units that can be maintained in WM is limited (Brady et al., 2009; Wood, 2007). Therefore, event models could be temporally compressed when the number of perceptual units they contain exceeds WM capacity. In general, the number of perceptual units increases with

---

[5] So far, we have used the term "event" to refer to meaningful units or "conceptual events" as it were. Perceptual units refer to a finer/shorter-scale segmentation (from a temporal point of view). Just as conceptual events correspond to periods during which the evolution of the current situation is highly predictable (Baldwin & Kosie, 2021), perceptual units correspond to periods during which sensorial inputs are highly predictable (Wood, 2007). Perceptual units are the building blocks of the conceptual events identified by human observers, and are hierarchically nested within them (Baldassano et al., 2017; Geerligs et al., 2021). Thus, changes delimiting perceptual units do not necessarily affect the meaning of the current situation. It is likely that these changes do not trigger a global update of the event model leading to its integration into long-term memory. Rather, they would induce incremental updates participating in the information accumulation process underlying event model construction (Lu et al., 2022; Richmond & Zacks, 2017).

event duration, but this is not always the case, and it would be interesting in future studies to investigate the independent effects of each factor on temporal compression. For that purpose, the perceptual units contained in video segments could be identified using behavioral data or an algorithmic approach (Rui & Anandan, 2000; Zacks & Tversky, 2001).[6]

In our view, the fact that RD starts to be shorter than the actual event duration for events lasting more than 9 s reflects a temporal limit of WM in representing continuous events. Beyond that duration, the stream of experience accumulated and integrated into the current event model would exceed WM capacity. As a result, the mental representation of the event would start to be truncated (i.e., part of the sensory stream is no longer represented), leading to an incomplete encoding of its unfolding in episodic memory. However, it could also be the case that some form of compression mechanism already operates at the level of WM representations (i.e., when constructing the event model of the current situation). Between event boundaries, the continuous unfolding of experience could be summarized through WM processes, enabling the construction of event models covering longer time spans, but in a format further away from reality/perception (Al Roumi et al., 2021; Brady et al., 2009; Lazartigues et al., 2021). Indeed, previous studies have shown that visuospatial information (e.g., spatial sequences, multidimensional objects, or color-changing shapes) that includes regularities tends to be compressed in WM. Compressed

---

[6] It should also be noted that event boundaries can be endogenously generated (i.e., without any meaningful change in perceptual input). Indeed, a recent study by Jafarpour et al. (2022) suggests that new event boundaries could be created when WM is overloaded. Furthermore, event boundaries can be generated following changes in affective, goal, and motivational states (Wang et al., 2023).Such endogenously generated event boundaries might influence how events are memorized despite little changes in perceptual information.

representations could be formed in WM by removing redundant information, thus allowing more items to be stored (Brady et al., 2009). Although this remains to be investigated, dynamic events containing important regularities in their unfolding (such as those depicted in our videos) could be subject to similar WM compression processes when they exceed a given duration.

The current finding that continuous events are temporally compressed in memory when they exceed a given duration could be due not only to factors operating during encoding, but also during retrieval. Event models are multidimensional and temporally extended representations that integrate the different elements characterizing the ongoing situation (e.g., people, locations, objects, and so on), as well as their temporal evolution (Bird, 2020; Loschky et al., 2020; Richmond & Zacks, 2017; Zacks et al., 2007). The retrieval of this within-event temporal evolution (the quantity of within-event information that can be retrieved and maintained online for the purpose of mental replay) could be limited by WM capacity. As a result, the mental replay of long continuous events could be truncated. Although this is a challenging undertaking, future studies should attempt to clarify the respective contribution of encoding and retrieval processes in determining RD, and thus temporal compression, for long continuous events.

Interestingly, we found that the estimated RD for short events (i.e., 3- and 6-s events) was slightly longer than the actual event duration (temporal ratio > 1). In fact, it is likely that our measure of RD (asking participants to press a key when they started and finished their mental replay) reflects the combined durations of two distinct mental processes: the time needed to mentally replay the unfolding of the event (our process of interest) and the time taken to access the initial memory trace (i.e., the time needed to represent the initial visual scene from which the participants had to mentally replay the subsequent unfolding of the event). Indeed, we can realistically assume that a minimal amount of time is required to initiate memory retrieval (McElree & Dosher, 1989; Staresina & Wimber, 2019). When this is taken into account, the

temporal ratio would necessarily be higher than 1 for events that are mentally replayed without temporal compression. Furthermore, the temporal ratio would be higher for 3- than 6-s events: for example, if someone needs 500 ms to access the initial visual scene, this represents 1/6 of the event duration for 3-s events (temporal ratio = 1.17) but only 1/12 of the event duration for 6-s events (temporal ratio = 1.08). Therefore, it is possible that our measure somewhat overestimated mental replay duration and, therefore, WM temporal limit (i.e., the duration beyond which, in the absence of event boundaries, temporal compression emerges). In addition, the time needed to access the initial visual scene might increase with event duration; consequently, our estimates of temporal compression might be particularly underestimated for the longest events. Further studies controlling for the time needed to access the initial memory trace would be necessary to quantify the temporal limit of WM more precisely.

Another issue that deserves further discussion is to what extent the present results relate to previous findings on memory for durations. Indeed, it could be that mental replay and duration estimation rely in part on common cognitive processes[7]. Remembered duration is constructed, at least in part, from the memory representation of "what happened" during the targeted time interval (i.e., from non-temporal information; Block & Reed, 1978; Block & Zakay, 1997). In line with this view, estimated durations for events (of the same length) tend to increase with the number of event boundaries (Faber & Gennari, 2015; Roseboom et al., 2022; Wang & Gennari,

---

[7] This could explain the similarity between the decrease of temporal ratio with event duration observed in the current study and the decrease of the "estimated duration/actual duration" ratio typically observed with tasks assessing memory for durations (e.g., Faber & Gennari, 2015; Gümüş & Balcı, 2023; Herbst et al., 2021; Roseboom et al., 2022; Wang & Gennari, 2019).

2019)[8]. Interestingly, this effect is thought to result from the weak encoding of within-event information (moments temporally distant from event boundaries; Wang & Gennari, 2019). Thus, although the mental replay of past events (as described in our task, at least) focuses on the mental reproduction of visuospatial information evolving over time rather than on duration as such, memory for events and durations present some overlap since they partly rely on the same (sometimes incomplete) source of information (i.e., past event models).

In other respects, memory for the duration of time intervals that do not include perceptual changes (e.g., "empty" time interval delimited by two auditory tones; Herbst et al., 2021) is thought to involve the memorization of temporal information itself (through the allocation of attentional resources to track the passage of time; Block & Zakay, 1997, 1997). Seminal theories proposed the idea of a mental pacemaker, producing pulses at a particular rate according to the organism's arousal. Pulses captured by attention would be accumulated in WM to form mental representations of duration. As a result, when available WM/attentional resources are reduced (e.g., because of a concurrent task; Block & Zakay, 1997; Dutke, 2005), the number of pulses integrated in the WM representation of duration would be reduced, leading to gaps in this representation and thus to shortened duration estimations (Gibbon et al., 1984). Therefore, while the object of remembering is different, WM capacity limit in accumulating continuous

---

[8] Note that we are talking here about low-level event boundaries detectable from stability of activity patterns in low-level sensory regions (Baldassano et al., 2017; Fountas et al., 2022; Geerligs et al., 2021), the kind of event boundary delimiting basic actions properties such as motion (the short perceptual events we mentioned earlier).

information could be one feature of our cognitive system explaining similarities between duration reproduction of changeless time intervals and the temporal aspect of event mental replay.

5. Conclusion

In conclusion, this study sheds new light on the cognitive mechanisms determining the formation of experience units representing past events in memory. We showed that continuous events begin to be temporally compressed in WM when they exceed a duration of about 9 s. This temporal compression of events could be a byproduct of a temporal WM limit (i.e., WM capacity limit in representing the continuous stream of sensory input) or could reflect the operation of compression mechanisms that are already at play when building event models in WM. These findings and our novel paradigm investigating WM for continuous events open new avenues of research to determine the precise duration of WM temporal limit and its contribution to the formation and retrieval of memories for real-world events.

References

Al Roumi, F., Marti, S., Wang, L., Amalric, M., & Dehaene, S. (2021). Mental compression of spatial sequences in human working memory using numerical and geometrical primitives. *Neuron*, *109*(16), 2627-2639.e4. https://doi.org/10.1016/j.neuron.2021.06.009

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Bailey, H. R., Kurby, C. A., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory & Cognition*, *45*(6), 940-955. https://doi.org/10.3758/s13421-017-0707-2

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10.1016/j.neuron.2017.06.041

Baldwin, D. A., & Kosie, J. E. (2021). How Does the Mind Render Streaming Experience as Events? *Topics in Cognitive Science*, *13*(1), 79-105. https://doi.org/10.1111/tops.12502

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83-100. https://doi.org/10.1037/0096-3445.133.1.83

Bird, C. M. (2020). How do we remember events? *Current Opinion in Behavioral Sciences*, *32*, 120-125. https://doi.org/10.1016/j.cobeha.2020.01.020

Block, R. A., & Reed, M. A. (1978). Remembered Duration : Evidence for a Contextual-Change Hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, *4*(6), 656-665. https://doi.org/10.1037/0278-7393.4.6.656

Block, R. A., & Zakay, D. (1997). Prospective and retrospective duration judgments : A meta-analytic review. *Psychonomic Bulletin & Review*, *4*(2), 184-197. https://doi.org/10.3758/BF03209393

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory : Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487-502. https://doi.org/10.1037/a0016797

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models : A Tutorial. *Journal of Cognition*, *1*(1), 9. https://doi.org/10.5334/joc.10

Conway, M. A. (2009). Episodic memories. *Neuropsychologia*, *47*(11), 2305-2313. https://doi.org/10.1016/j.neuropsychologia.2009.02.003

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognition and Development*, *11*(2), 121-136. https://doi.org/10.1080/15248371003699969

D'Argembeau, A. (2020). Zooming In and Out on One's Life : Autobiographical Representations at Multiple Time Scales. *Journal of Cognitive Neuroscience*, *32*(11), 2037-2055. https://doi.org/10.1162/jocn_a_01556

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. https://doi.org/10.1080/09658211.2021.1896737

DeBruine, L. M., & Barr, D. J. (2021). Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1-15. https://doi.org/10.1177/2515245920965119

Dutke, S. (2005). Remembered duration : Working memory and the reproduction of intervals. *Perception & Psychophysics*, *67*(8), 1404-1413. https://doi.org/10.3758/BF03193645

Eisenberg, M. L., & Zacks, J. M. (2016). Ambient and focal visual processing of naturalistic activity. *Journal of Vision*, *16*(2), 5. https://doi.org/10.1167/16.2.5

Faber, M., & Gennari, S. P. (2015). In search of lost time : Reconstructing the unfolding of events from memory. *Cognition*, *143*, 193-202. https://doi.org/10.1016/j.cognition.2015.06.014

Fairhall, S. L., Albi, A., & Melcher, D. (2014). Temporal Integration Windows for Naturalistic Visual Sequences. *PLoS ONE*, *9*(7), e102248. https://doi.org/10.1371/journal.pone.0102248

Fountas, Z., Sylaidi, A., Nikiforou, K., Seth, A. K., Shanahan, M., & Roseboom, W. (2022). A Predictive Processing Model of Episodic Memory and Time Perception. *Neural Computation*, *34*(7), 1501-1544. https://doi.org/10.1162/neco_a_01514

Geerligs, L., Van Gerven, M., & Güçlü, U. (2021). Detecting neural state transitions underlying event segmentation. *NeuroImage*, *236*, 118085. https://doi.org/10.1016/j.neuroimage.2021.118085

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar Timing in Memory. *Annals of the New York Academy of Sciences*, *423*(1 Timing and Ti), 52-77. https://doi.org/10.1111/j.1749-6632.1984.tb23417.x

Güler, B., Adıgüzel, Z., Uysal, B., & Gunseli, E. (2023). *Discrete memories of a continuous world : A working memory perspective on event segmentation* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/4mx79

Gümüş, G., & Balcı, F. (2023). Working memory for time intervals : Another manifestation of the central tendency effect. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02324-z

Herbst, S. K., Mangione, I., Kononowicz, T., & van Wassenhove, V. (2021). *Abstracting time in memory* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/478cx

Jafarpour, A., Buffalo, E. A., Knight, R. T., & Collins, A. G. E. (2022). Event segmentation reveals working memory forgetting rate. *iScience*, *25*(3), 103902. https://doi.org/10.1016/j.isci.2022.103902

Jeneson, A., & Squire, L. R. (2012). Working memory, long-term memory, and medial temporal lobe function. *Learning & Memory*, *19*(1), 15-25. https://doi.org/10.1101/lm.024018.111

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. https://doi.org/10.1177/1747021818773082

Jeunehomme, O., & D'Argembeau, A. (2020). Event segmentation and the temporal compression of experience in episodic memory. *Psychological Research*, *84*(2), 481-490. https://doi.org/10.1007/s00426-018-1047-y

Jeunehomme, O., & D'Argembeau, A. (2023). Memory editing : The role of temporal discontinuities in the compression of events in episodic memory editing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(5), 766-775. https://doi.org/10.1037/xlm0001141

Koller, M. (2013). *Robust Estimation of Linear Mixed Models* [Doctoral thesis, ETH Zurich]. ETH Repository for Publications and Research Data. https://doi.org/10.3929/ethz-a-007632241

Koller, M. (2016). robustlmm : An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, *75*(6), 1-24. https://doi.org/10.18637/jss.v075.i06

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72-79. https://doi.org/10.1016/j.tics.2007.11.004

Kurby, C. A., & Zacks, J. M. (2011). Age differences in the perception of hierarchical structure in events. *Memory & Cognition*, *39*(1), 75-91. https://doi.org/10.3758/s13421-010-0027-2

Lazartigues, L., Lavigne, F., Aguilar, C., Cowan, N., & Mathy, F. (2021). Benefits and pitfalls of data compression in visual working memory. *Attention, Perception, & Psychophysics*, *83*(7), 2843-2864. https://doi.org/10.3758/s13414-021-02333-x

Loschky, L. C., Larson, A. M., Smith, T. J., & Magliano, J. P. (2020). The Scene Perception & Event Comprehension Theory (SPECT) Applied to Visual Narratives. *Topics in Cognitive Science*, *12*(1), 311-351. https://doi.org/10.1111/tops.12455

Lu, Q., Hasson, U., & Norman, K. A. (2022). A neural network model of when to retrieve and encode episodic memories. *eLife*, *11*, e74445. https://doi.org/10.7554/eLife.74445

Mason, F. (2022). *Inférence Statistique dans les Modèles Linéaires Mixtes avec estimation robuste, appliqués aux données longitudinales* [Doctoral thesis, Geneva University]. Open archive UNIGE. https://archive-ouverte.unige.ch/unige:161871

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305-315. https://doi.org/10.1016/j.jml.2017.01.001

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory : The time course of recognition. *Journal of Experimental Psychology: General*, *118*(4), 346-373. https://doi.org/10.1037/0096-3445.118.4.346

Mirman, D. (2014). *Growth curve analysis and visualization using R* (1st éd.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315373218

Monfort, M., Vondrick, C., Oliva, A., Andonian, A., Zhou, B., Ramakrishnan, K., Bargal, S. A., Yan, T., Brown, L., Fan, Q., & Gutfreund, D. (2020). Moments in Time Dataset : One Million Videos for Event Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(2), 502-508. https://doi.org/10.1109/TPAMI.2019.2901464

Montemayor, C., & Wittmann, M. (2014). The Varieties of Presence : Hierarchical Levels of Temporal Integration. *Timing & Time Perception*, *2*(3), 325-338. https://doi.org/10.1163/22134468-00002030

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, *1*(2), 56-61. https://doi.org/10.1016/S1364-6613(97)01008-5

Posit team. (2023). *RStudio : Integrated development environment for R* [Logiciel]. Posit Software, PBC. http://www.posit.co/

R Core Team. (2022). *R: A language and environment for statistical computing* [Logiciel]. R Foundation for Statistical Computing. https://www.R-project.org/

Radvansky, G. A. (2017). Event segmentation as a working memory process. *Journal of Applied Research in Memory and Cognition*, *6*(2), 121-123. https://doi.org/10.1016/j.jarmac.2017.01.002

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, *17*, 133-140. https://doi.org/10.1016/j.cobeha.2017.08.006

Richmond, L. L., & Zacks, J. M. (2017). Constructing Experience : Event Models from Perception to Action. *Trends in Cognitive Sciences*, *21*(12), 962-980. https://doi.org/10.1016/j.tics.2017.08.005

Roseboom, W., Seth, A., Sherman, M. T., & Fountas, Z. (2022). *The Perception of Time in Humans, Brains, and Machines* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/c7vzx

Rui, Y., & Anandan, P. (2000). Segmenting visual actions based on spatio-temporal motion patterns. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, *1*, 111-118. https://doi.org/10.1109/CVPR.2000.855807

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., Eisenberg, M. L., & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*(2), 241-255. https://doi.org/10.1016/j.cognition.2013.07.002

Smith, M. E., Loschky, L. C., & Bailey, H. R. (2021). Knowledge guides attention to goal-relevant information in older adults. *Cognitive Research: Principles and Implications*, *6*(1), 56. https://doi.org/10.1186/s41235-021-00321-1

Smith, M. E., Newberry, K. M., & Bailey, H. R. (2020). Differential effects of knowledge and aging on the encoding and retrieval of everyday activities. *Cognition*, *196*, 104159. https://doi.org/10.1016/j.cognition.2019.104159

Staresina, B. P., & Wimber, M. (2019). A Neural Chronometry of Memory Recall. *Trends in Cognitive Sciences*, *23*(12), 1071-1085. https://doi.org/10.1016/j.tics.2019.09.011

Tulving, E. (2002). Episodic Memory : From Mind to Brain. *Annual Review of Psychology*, *53*(1), 1-25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Wahlheim, C. N., Eisenberg, M. L., Stawarczyk, D., & Zacks, J. M. (2022). Understanding Everyday Events : Predictive-Looking Errors Drive Memory Updating. *Psychological Science*, *33*(5), 765-781. https://doi.org/10.1177/09567976211053596

Wang, Y. C., Adcock, R. A., & Egner, T. (2023). Toward an integrative account of internal and external determinants of event segmentation. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02375-2

Wang, Y., & Gennari, S. P. (2019). How language and event recall can shape memory for time. *Cognitive Psychology*, *108*, 1-21. https://doi.org/10.1016/j.cogpsych.2018.10.003

Wittmann, M. (2016). *Felt time : The psychology of how we perceive time* (E. Butler, Trad.). MIT Press.

Wood, J. N. (2007). Visual working memory for observed actions. *Journal of Experimental Psychology: General*, *136*(4), 639-652. https://doi.org/10.1037/0096-3445.136.4.639

Zacks, J. M. (2020). Event Perception and Memory. *Annual Review of Psychology*, *71*(1), 165-191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception : A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273-293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3-21. https://doi.org/10.1037/0033-2909.127.1.3

Supplementary Materials

Supplementary information regarding statistical analyses

R packages

Data cleaning, formatting and pre-processing were conducted with the packages dplyr (version 1.1.2; Wickham et al., 2023) and purr (version 0.3.5; Wickham & Henry, 2023). The R project was built with renv (version 0.17.3; Ushey & Wickham, 2023). Tables were made with the packages flextable (version 0.9.1; Gohel & Skintzos, 2023) and rempsyc (version 0.1.1; Thériault, 2022). Figures were made with the packages ggplot2 (version 3.4.0; Wickham, 2016) and ggpubr (version 0.4.0; Kassambara, 2020). The original draft of this article was written using rmarkdown (version 2.21; Xie et al., 2020), knitr (version 1.42; Xie, 2023) and pandoc (version 3.1.2; Mailund, 2019). Application conditions of classical linear mixed-effects models were checked with the performance package (version 0.10.3; Lüdecke et al., 2021). The robust estimation procedure (the DAStau procedure) was implemented using the robustlmm package (version 3.0.4; Koller, 2016). We relied on functions from the packages broom.mixed (version 0.2.9.4; Bolker & Robinson, 2022) and parameters (version 0.21.0; Lüdecke et al., 2020) to compute confidence intervals (*CI*s) and *p*-values (for fixed effects) from standard errors (*SEs*) and *t*-statistics. Estimated remembering duration (RD) and temporal ratio, associated pointwise standard errors and Wald's 95% CIs were computed for each event duration using variance-covariance matrix of the models with the package effects (version 4.2.2; Fox & Weisberg, 2019).

Power analysis

The targeted sample size for this study was determined a priori with a power analysis based on Monte-Carlo simulations (Brysbaert & Stevens, 2018; DeBruine & Barr, 2021). We conducted a pilot study with 10 participants who were not included in our final sample (data and scripts from this pilot study and power analyses are available at https://osf.io/zkwf2). Using these pilot data, we fitted each model we plan to run on our final dataset. Then, we computed a specific power curve (i.e., the statistical power that can be achieved with a range of sample sizes) for each of these models. The alpha level was set to 0.05. Taken together, these analyses indicated that a sample size of 90 participants would provide us with a statistical power of at least 90% for each of the planned analyses.

Data cleaning

As preregistered, participants who did not complete the entire experimental task were excluded from the analyses and experimental trials for which there was a technical issue with the video presentation (e.g., an abnormal duration of presentation) were excluded. In addition, we removed trials with a time of mental replay shorter than 1 s (i.e., to remove trials during which the participant inadvertently pressed the button twice or did not properly follow the instructions) or longer than twice the real duration of the remembered video (i.e., to remove trials during which the participant was interrupted or did not mentally replay the video in one shot). If more than half of the trials of one type (one duration) had to be excluded for a given participant, the entire participant was removed from the analysis. Each removed participant was replaced by another. In total, 17 participants were excluded: seven because the demographic questionnaire filled at the end of the experiment revealed that they did not match the inclusion criteria and ten because they did not provide enough valid trials (according to the criteria mentioned above).

Model's performance

We evaluated the quality of the model predicting temporal ratios (see Main Text) with five indices (computed with the package performance; version 0.10.3; Lüdecke et al., 2021). First, the marginal and conditional Nakagawa's R-squared (R^2s). Marginal R^2 represents the part of the dependent variable variance explained by fixed effects alone, while conditional R^2 represents the variance explained by the entire model (fixed and random effects; Johnson, 2014; Nakagawa et al., 2017; Nakagawa & Schielzeth, 2013). The part of variance explained by the grouping structure of our data was quantified with intra-class correlation (*ICC*; Hox, 2010; Rabe-Hesketh & Skrondal, 2012; Raudenbush & Bryk, 2002). The accuracy and precision of models' predicted values were assessed using the Root Mean Square Error (*RMSE;* Chai & Draxler, 2014; Kenney & Keeping, 1962) and the residual standard deviation (*sigma*; $\boldsymbol{\sigma}$; Gelman et al., 2020).

The model explained 89% of the variance in temporal ratio (Conditional $R^2 = 0.89$) and its fixed part 35% (Marginal $R^2 = 0.35$). Notably, the grouping structure in our data (several temporal ratios measured for each participant) explained a substantial part of variance (*ICC* = 0.84). Accuracy and precision of the model were satisfying. Considering that actual temporal ratios ranged between 0.2 and 1.73, the average distance between data and the model trendline was quite low (*RMSE* = 0.06), as was the average gap between fitted and observed values ($\boldsymbol{\sigma}$ = 0.08).

Main analyses with classical linear mixed-effects models (non-robust estimations)

Following our pre-registered analyses plan, we tested our hypotheses using linear-mixed effects models (lmm) and used robust estimations as the assumptions underlying classical statistical analyses were not fully met (see https://osf.io/zkwf2 for the complete assessment). However, for the sake of completeness, we here report results of classical linear mixed-effects models fitted with the package lme4 (maximum likelihood, nloptwrap optimizer; version 1.1.31; Bates et al., 2015). Inferential tests on fixed effects were conducted with the package lmerTest (version 3.1.3; Kuznetsova et al., 2017) using Satterthwaite approximation for degrees of freedom (Luke, 2017). The pattern of results was similar to the robust analyses reported in the main text.

Remembering duration (RD) as a function of event duration

For each participant, RDs were averaged by stimulus duration and we examined the relation between stimulus duration and the time taken by participants to mentally replay the videos by fitting a model (growth curve analysis) with RD as outcome and a first and a second order polynomial transformation of stimulus duration as predictors. This model included two uncorrelated random effects: a random intercept for participants and a random slope for the linear term. As its robust counterpart (reported in the main text), this first model revealed a non-linear increase of RD as a function of event duration (see Table S1). To further check the statistical significance of the quadratic term, we also compared the model including both linear and quadratic terms to a simpler model that only included the linear term (using a likelihood ratio test, following the model comparison approach proposed by Mirman (2014). This comparison was performed using the "anova" function of the R Stats package (version 4.2.2; R Core Team, 2022) and showed that the quadratic term significantly increased the model fit ($\chi^2(1) = 75.69$, $p < 0.01$).

Bootstrap conditional 95% *CI*s were computed for each fitted value with the package ciTools

(version 0.6.1; Haman & Avery, 2020; see Figure S1, Table S3). Importantly, RD was close to

the actual event duration for short events (i.e., 3 and 6-s events) but significantly shorter for long

events (12 and 15-s events).

Table S1

Growth curve analysis: prediction of RD by event duration (non-robust)

Term	*b*	*SE*	*t*	*df*	*p*	95% CI
Intercept	8.37	0.19	42.99	90.00	< .001	[7.99, 8.76]
Linear term	6.71	0.27	24.96	90.00	< .001	[6.18, 7.24]
Quadratic term	-0.73	0.08	-9.35	270.00	< .001	[-0.89, -0.58]

Note. Coefficients of fixed effects and related inferential statistics.

Temporal ratio as a function of event duration

Next, we examined the relation between stimuli duration and temporal ratio by fitting a second growth curve with temporal ratio as outcome and a first and a second order polynomial transformation of stimuli duration as predictors. This model included three uncorrelated random effects: a random intercept for participants, and random slopes for the linear and quadratic terms. As expected we observed a non-linear decrease of temporal ratio as a function of event duration (see Figure S1, Table S2) and the quadratic term significantly increased the model fit ($\chi^2(1) =$ 12.33, p <0.01). The model estimates showed that mental replay durations became significantly shorter than the actual event durations when events lasted more than 9 s (i.e., temporal ratios were significantly lower than 1 only for 12- and 15-s events; see Table S3).

Table S2

Growth curve analysis: prediction of temporal ratio by event duration (non-robust)

Term	b	SE	t	df	p	95% CI
(Intercept)	1.00	0.02	48.82	90.00	< .001	[0.96, 1.04]
Linear term	-0.32	0.02	-15.73	90.00	< .001	[-0.37, -0.28]
Quadratic term	0.04	0.01	3.64	90.00	< .001	[0.02, 0.07]

Note. Model performance indexes: *R2_conditional* = 0.89, *R2_marginal* = 0.33, *ICC* = 0.84, *RMSE* = 0.06, *Sigma* = 0.08.

Figure S1

Evolution of RD and temporal ratio as a function of event duration (non-robust)
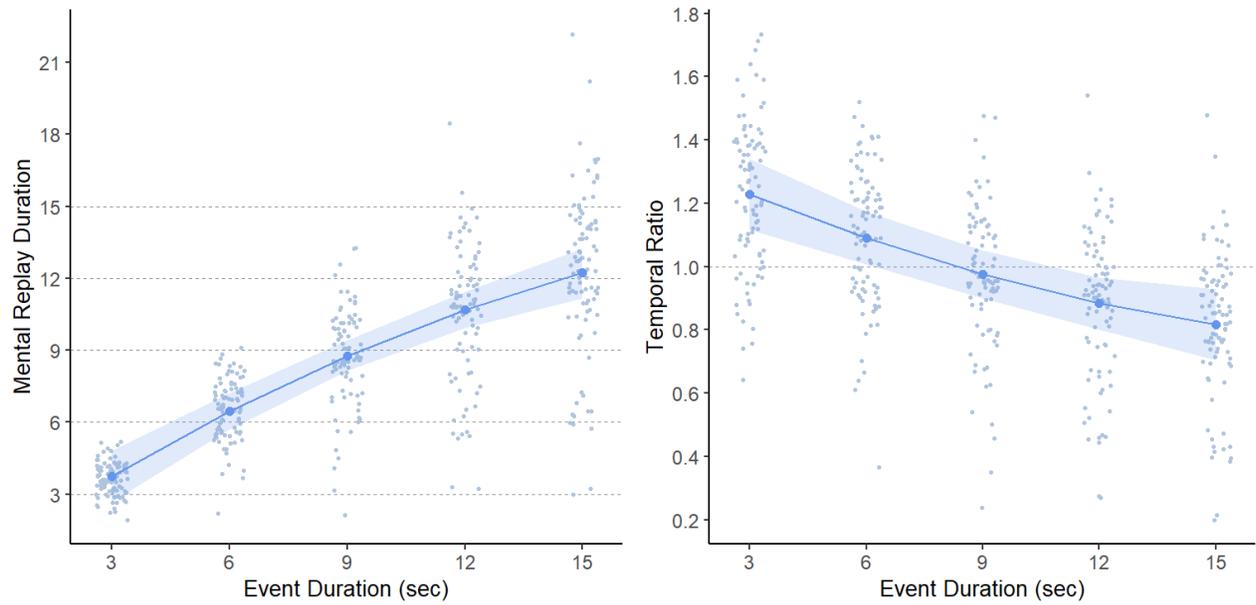
Note. The blue-gray dots represent the observed values. Each point corresponds to the average value of a given participant for a given duration. The dark-blue dots and the ribbon surrounding them represent the model estimates and their 95% CIs for each event duration.

Table S3

Growth curve analysis: estimates across event durations (non-robust)

Event Duration	Remembering Duration		Temporal Ratio	
	Estimate	95%CI	Estimate	95%CI
3	3.74	[2.69, 4.80]	1.23	[1.12, 1.34]
6	6.45	[5.69, 7.21]	1.09	[1.01, 1.17]
9	8.77	[8.13, 9.40]	0.97	[0.90, 1.05]
12	10.69	[9.94, 11.44]	0.88	[0.80, 0.97]
15	12.22	[11.18, 13.25]	0.82	[0.71, 0.93]

Note. Estimated values and bootstrap 95% *CI*s for the five event durations.

Main analyses with non-cleaned data

To ensure that the results we reported in the Main Text were not dependent on our inclusion and data cleaning criteria, we fitted the same two robust linear mixed-effects models as described in the Main Text on the non-cleaned data (N participants = 107, N observations = 2675). Results of these supplementary analyses are presented below.

RD as a function of event duration

RD non-linearly increased with event duration (see Table S4, Figure S2). The increase in RD was lower as event duration increased (e.g., the increase in RD between 3 and 6-s events was greater than the one between 12 and 15-s events; see Table S5).

Table S4

Growth curve analysis: prediction of RD by event duration (non-cleaned data)

Term	*b*	*SE*	*t*	95% CI
Intercept	8.76	0.19	46.48	[8.39, 9.13]
Linear Term	6.74	0.26	25.50	[6.22, 7.26]
Quadratic Term	-0.68	0.07	-9.26	[-0.82, -0.53]

Note. Coefficients of fixed effects and related inferential statistics

Table S5

Growth curve analysis: estimates across event durations (non-cleaned data)

Event Duration	Remembering Duration		Temporal Ratio	
	Estimate	95%CI	Estimate	95%CI
3	3.74	[2.69, 4.80]	1.23	[1.12, 1.34]
6	6.45	[5.69, 7.21]	1.09	[1.01, 1.17]
9	8.77	[8.13, 9.40]	0.97	[0.90, 1.05]
12	10.69	[9.94, 11.44]	0.88	[0.80, 0.97]
15	12.22	[11.18, 13.25]	0.82	[0.71, 0.93]

Note. Estimated values and bootstrap 95% *CI*s for the five event durations.

Temporal ratio as a function of event duration

Similarly, a growth curve analysis with temporal ratio as outcome and a first and a second order polynomial transformation of stimulus duration as predictors revealed a non-linear decrease of temporal ratio with event duration (both the linear and the quadratic terms were statistically significant; see Table S6). More specifically, temporal ratio was significantly smaller than 1 (reflecting a mental replay duration inferior to the actual video duration) only for 12- and 15-s events (Table S5 and Figure S2).

Table S6

Growth curve analysis: prediction of temporal ratio by event duration (non-cleaned data)

Term	*b*	*SE*	*t*	95% CI
(Intercept)	1.05	0.02	48.26	[1.01, 1.09]
Linear Term	-0.39	0.03	-14.60	[-0.44, -0.33]
Quadratic Term	0.08	0.01	5.62	[0.05, 0.11]

Note. Model performance indexes: *R2_conditional* = 0.92, *R2_marginal* = 0.37, *ICC* = 0.87, *RMSE* = 0.07, *Sigma* = 0.08.

Figure S2

*Evolution of RD and temporal ratio as a function of event duration (non-cleaned data)*
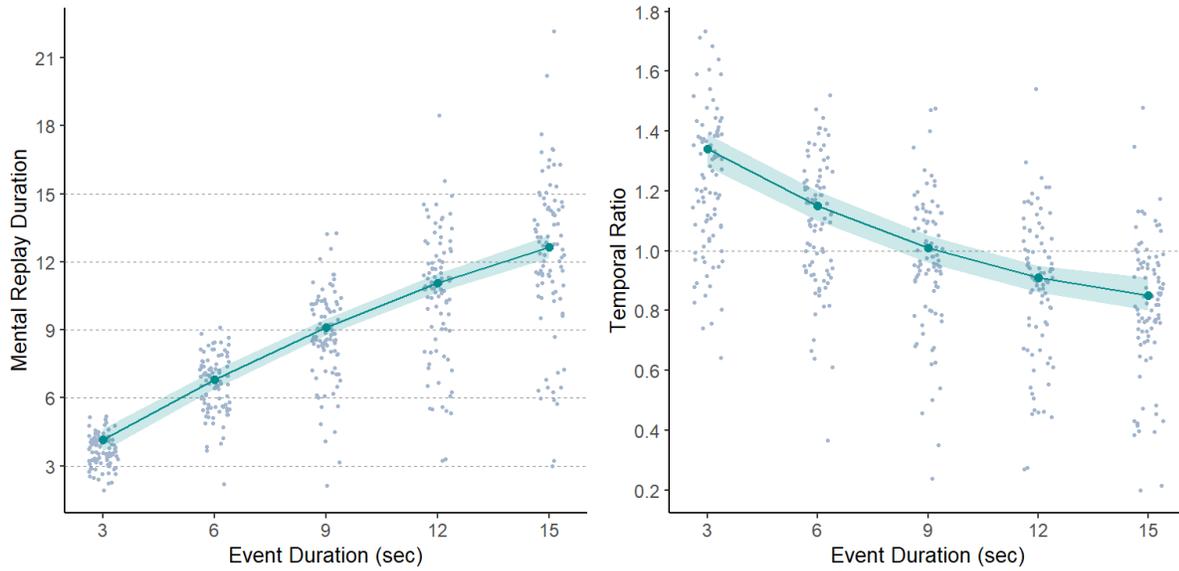
Note. The gray dots represent the observed values. Each point corresponds to the average remembering duration of a given participant for a given duration. The green dots and the ribbon surrounding them represent the model estimates and their 95% CIs for each event duration.

Decrease of temporal ratio as a function of event duration: individual slopes

According to the model estimates, the slopes of the linear and quadratic terms varied substantially in magnitude across participants. In other words, the extent to which the temporal ratio decreased with event duration, and the degree of curvature of this decrease, differed from one individual to another. Nevertheless, the direction of both linear and quadratic terms coefficients remained highly consistent (temporal ratios decreased non-linearly with event duration in most participants; see Figures S3 and S4).
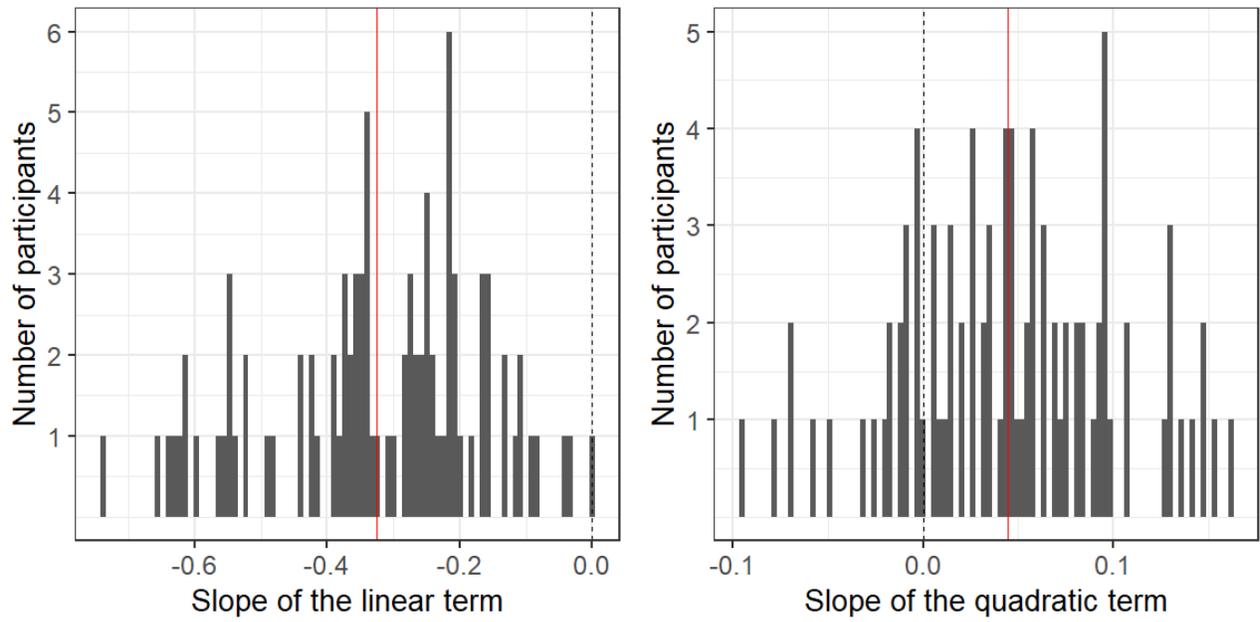
Figure S3

Decrease of temporal ratio as a function of event duration
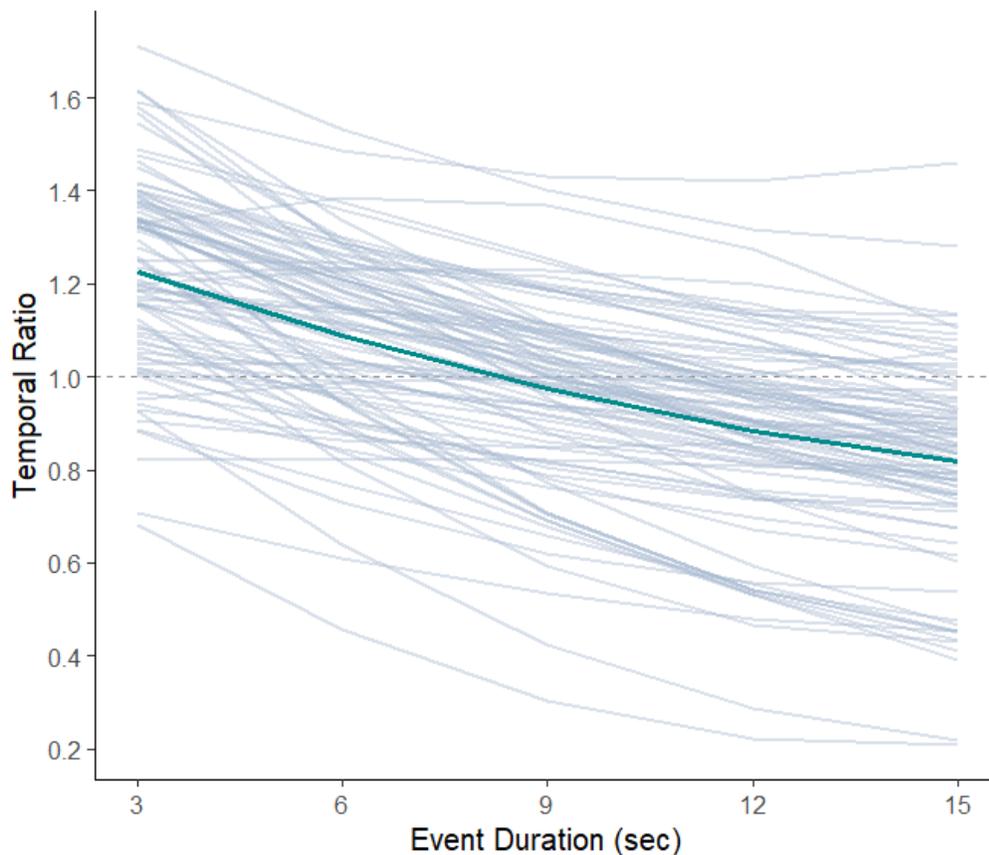
Note. Each grey line represents the individual slope of a participant. The green line represents the global trend according to the model estimates.

Figure S4

Individual coefficients of each participant

Note. Variation of the linear and quadratic terms slopes between participants. The red vertical line represents the global coefficient. Left panel: linear term. Right panel: quadratic term.

Descriptive statistics

The Results section of the main text focuses on estimates from the robust linear mixed-effects models (growth curve analyses) fitted on aggregated data. For the sake of completeness, we here provide descriptive statistics for RDs and temporal ratios (see Table S7), as well as plots showing their distribution (Figure S5).

Table S7

Descriptive statistics (non-aggregated data)

Event Duration	Remembering Duration					Temporal Ratio				
	Min	Q1	Median	Q3	Max	Min	Q1	Median	Q3	Max
3	1.09	2.92	3.67	4.37	5.93	0.36	0.97	1.22	1.46	1.98
6	1.45	5.58	6.59	7.60	11.24	0.24	0.93	1.10	1.27	1.87
9	1.02	7.18	8.89	10.30	17.58	0.11	0.80	0.99	1.14	1.95
12	1.37	8.48	10.74	12.88	21.85	0.11	0.71	0.89	1.07	1.82
15	2.10	9.77	12.50	14.92	27.18	0.14	0.65	0.83	0.99	1.81

Note. RD (left) and temporal ratio (right) as a function of event duration. From left to right: minimum, first quartile, median, third quartile and maximum.

Figure S5

Distribution of RD (left) and temporal ratio (right) for the five event durations (violin and box

plots)
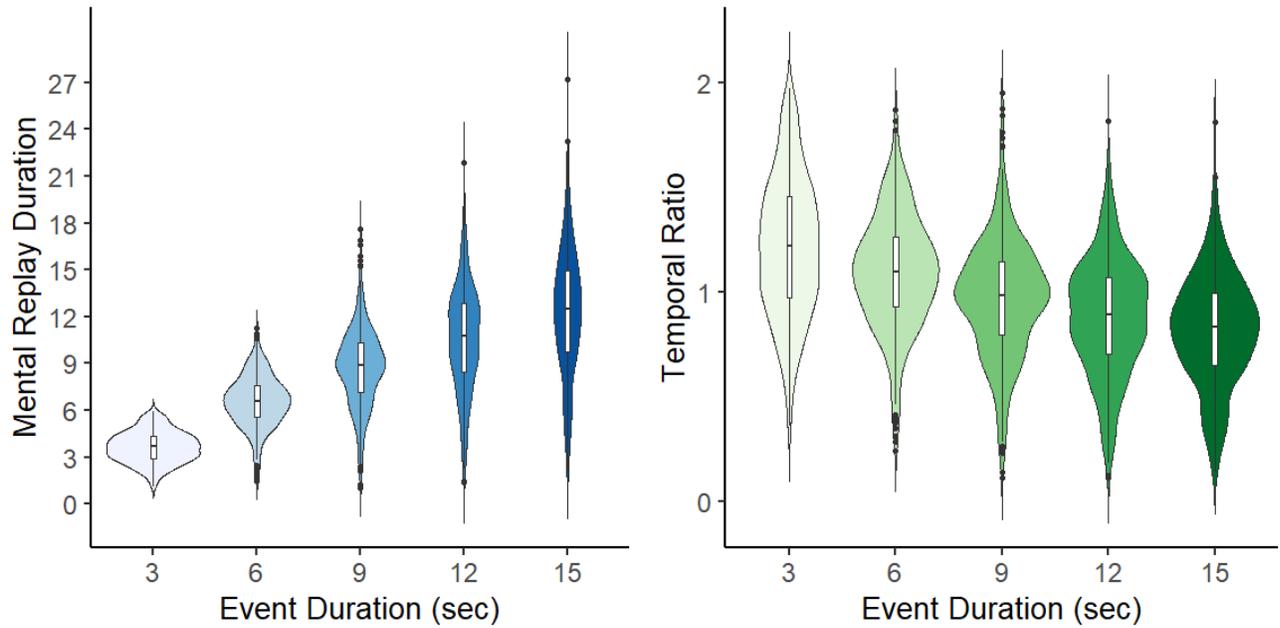
Note. Non-aggregated data.

Assessing the potential effect of central tendency bias in the relation between RD/temporal ratio and event duration

In line with our hypotheses, RD for long events (12 and 15-s) were shorter than the actual event duration (i.e., temporal ratio < 1). Unexpectedly, however, RD tended to be higher than event duration (temporal ratio > 1) for short events (3 and 6-s). In other words, temporal ratio was > 1 for events below the central event duration (i.e., 9 s) but < 1 for events above the central event duration. As a result, one could argue that the differences we observed between event duration and RD (i.e., variations in temporal ratio) could be explained by a central tendency bias rather than working memory capacity limit in representing continuous events. We conducted three supplementary analyses to assess this possibility.

Formal description of the central tendency bias goes back to Hollingworth (1910) who demonstrated it with a size judgment task (see Hellström, 1985, for a review of earlier reports of central tendency effects). It refers to the fact that estimated values in both reproduction and perceptual matching tasks tend to regress toward the mean of the distribution of stimuli magnitude (i.e., to be biased toward the center of the contextual distribution presented to participants; Aston et al., 2021; Duffy & Smith, 2018; Olkkonen et al., 2014)[9].

For the central tendency effect to take place, participants must be exposed to the entire distribution of stimuli magnitudes (i.e., to the different possible stimuli). Consequently, it should

---

[9] Although this effect was studied with a large range of psychological judgement/reproduction tasks involving various stimuli type (including time interval; e.g., Gümüş & Balcı, 2023; Jazayeri & Shadlen, 2010; Ryan, 2011), the nature of the cognitive processes driving it is still a matter of debate (Huttenlocher et al., 2000; Saarela et al., 2023; Wedell et al., 2020). However, further discussion of the different theoretical accounts of this bias would be beyond the scope of the current study.

not occur the first time participants are exposed to each possible stimulus. Thus, in the present study, if the same relation between event duration and temporal ratio (a non-linear decrease, with temporal ratios >1 for short durations and temporal ratios <1 for long durations) is still observed when participants were only exposed once to each duration, this behavioral effect could not be attributed to a central tendency bias.

Following this idea, we fitted the same growth curve models as those presented in the main text on a subset of our data including, for each participant, only the first trial of each event duration (five trials per participant, one per event duration). We obtained the same pattern of results as in the main analyses (a non-linear increase of RD as a function of event duration, with temporal ratio >1 for 3 and 6 s events and temporal ratio <1 for 12 and 15 s events; see Table S8 and Figure S6).

Table S8

Growth curve analysis: estimates across event durations (first five trials per participant, one per event duration)

Event Duration	Remembering Duration		Temporal Ratio	
	Estimate	95%CI	Estimate	95%CI
3	3.80	[3.25, 4.35]	1.23	[1.17, 1.28]
6	6.53	[6.11, 6.96]	1.10	[1.05, 1.14]
9	8.92	[8.51, 9.32]	0.99	[0.94, 1.04]
12	10.96	[10.53, 11.38]	0.91	[0.86, 0.95]
15	12.65	[12.09, 13.20]	0.85	[0.79, 0.91]

Figure S6

Evolution of RD (left) and temporal ratio (right) for the five event durations (first five trials per
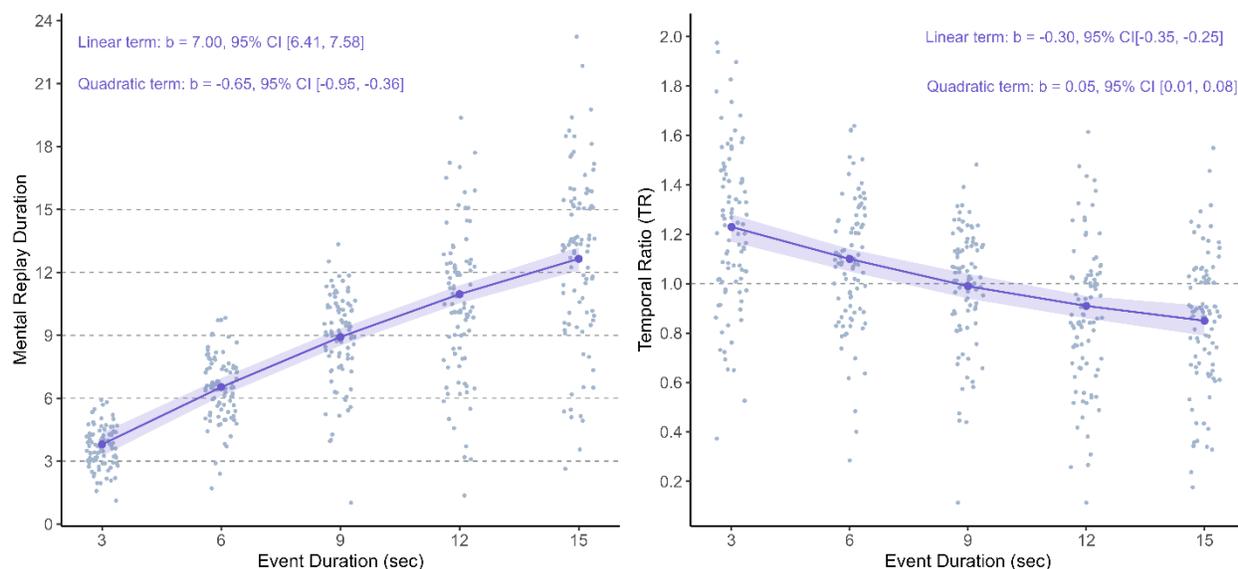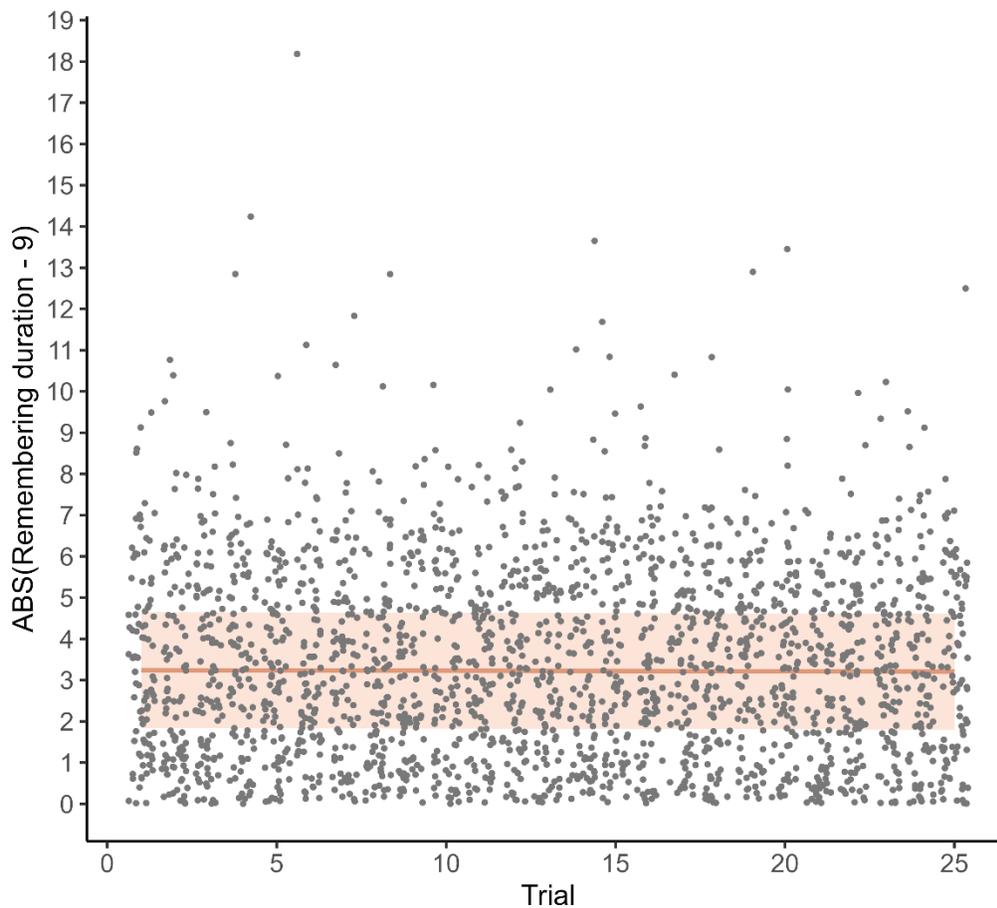
participant, one per event duration)

Note. The gray dots represent the observed values. Each point corresponds to the average remembering duration of a given participant for a given duration. The purple dots and the ribbon surrounding them represent the model estimates and their 95% CIs for each event duration.

Second, if a central tendency bias was at work, remembering durations should progressively approach 9 s (the center of the distribution of stimuli durations) when participants have been exposed several times to the range of possible durations (i.e., as the number of trials performed increases). To investigate this possibility, we fitted a robust linear mixed-effect model with the absolute value of the difference between RD and 9 s as outcome and the number of trials already performed as predictors (see Figure S7). This model included three random intercepts: one for participants, one for stimuli, and one for event duration. We did not observe any significant effect of the number of trials (Estimate = -0.001, *SE* = 0.005, 95% *CI* [-0.01, 0.01], *p*

= -0.776). Moreover, the variance explained the fixed part of the model was very small

(Marginal $R^2 = 0.00002$).

Figure S7

Deviation of RD from the center of the distribution of stimuli duration (9 s) across the trials

Note. The gray dots represent the observed values. Each point corresponds to the outcome of a single trial for a given participant. The orange line and the ribbon surrounding it represent the model estimates and their 95% CIs.

Finally, the fact that the differences between RD and event duration between short and long videos were asymmetrical (the difference between RD and median event duration was higher for 12-s and 15-s events compared to 3-s and 6-s events; see Table S7) also goes against

the view that the effect of event duration on temporal ratio we observed is entirely due to central

tendency bias. To inspect the statistical significance of this asymmetry, we fitted a robust linear

mixed effects model with the difference between RD and event duration as outcome and a first

and a second order polynomial transformation of stimulus duration as predictors (the model

included a random intercept for participants and random slopes the linear and quadratic terms).

According to the model estimates, the difference of interest was close to 0 for 3-, 6-, and 9-s

events but was more pronounced for 12- and 15-s events (linear term: b = -2.59, SE = 0.26,

95% CI [-3.10, -2.09], t = -10.10, p< .001; quadratic term: b = -0.74, SE = 0.08, 95% CI [-0.89, -

0.59], t = -9.71, p < .001; see Table S9 and figure S8). This asymmetry is hard to explain in terms
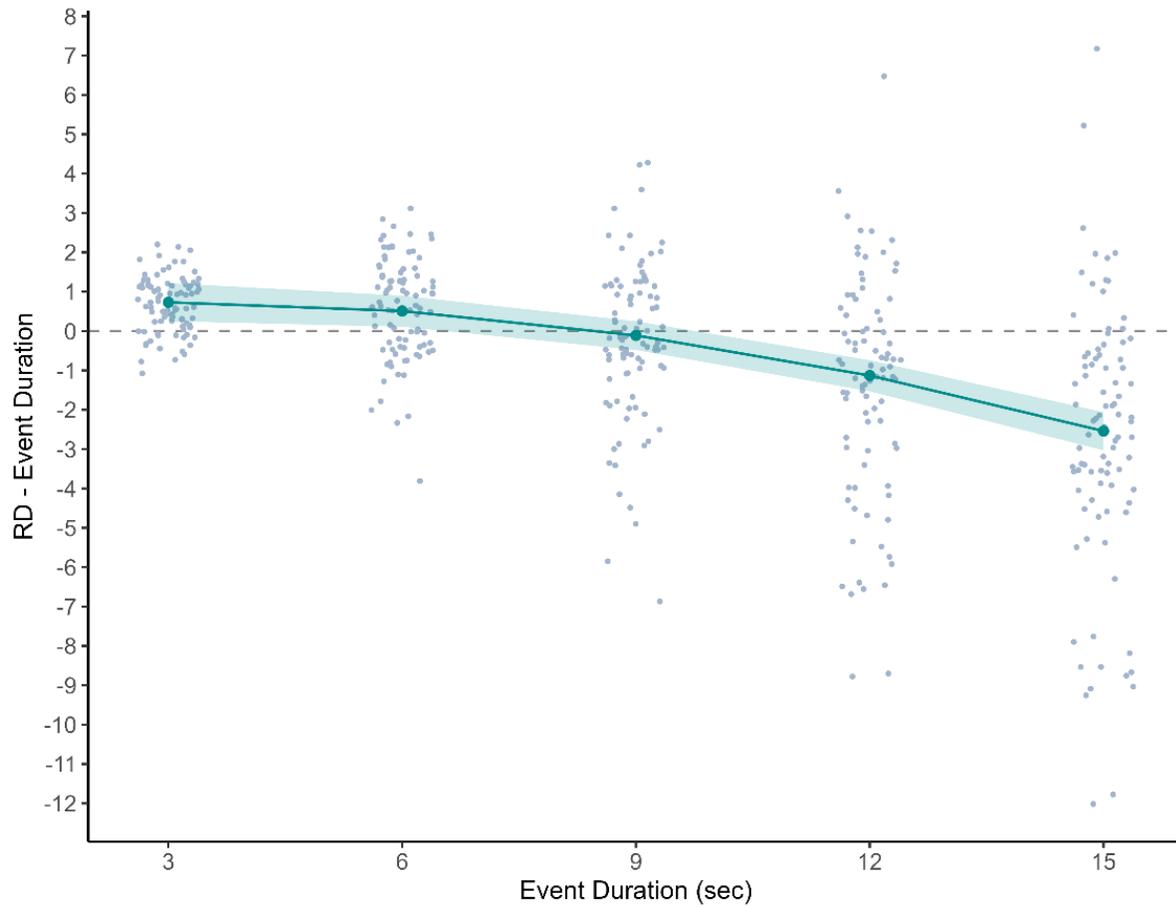
of central tendency effects alone.

Table S9

Difference between RD and event duration: estimates for each event duration

Duration	Estimate	95% CI
3 s	0.73	[0.25, 1.22]
6 s	0.51	[0.11, 0.90]
9 s	-0.11	[-0.48, 0.25]
12 s	-1.13	[-1.52, -0.74]
15 s	-2.54	[-3.03, -2.06]

Figure S8

Difference between RD and event duration across event durations

Note. The gray dots represent the observed values. Each point corresponds to the average remembering duration of a given participant for a given duration. The green dots and the ribbon surrounding them represent the model estimates and their 95% CIs for each event duration.

References

Aston, S., Negen, J., Nardini, M., & Beierholm, U. (2021). Central tendency biases must be accounted for to consistently capture Bayesian cue combination in continuous response data. *Behavior Research Methods*, *54*(1), 508-521. https://doi.org/10.3758/s13428-021-01633-2

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Bolker, B., & Robinson, D. (2022). *broom.mixed : Tidying methods for mixed models* [Logiciel]. https://CRAN.R-project.org/package=broom.mixed

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models : A Tutorial. *Journal of Cognition*, *1*(1), 9. https://doi.org/10.5334/joc.10

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014

DeBruine, L. M., & Barr, D. J. (2021). Understanding Mixed-Effects Models Through Data Simulation. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1-15. https://doi.org/10.1177/2515245920965119

Duffy, S., & Smith, J. (2018). Category effects on stimulus estimation : Shifting and skewed frequency distributions—A reexamination. *Psychonomic Bulletin & Review*, *25*(5), 1740-1750. https://doi.org/10.3758/s13423-017-1392-7

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd éd.). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/index.html

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and other stories*. Cambridge University Press. https://doi.org/10.1017/9781139161879

Gohel, D., & Skintzos, P. (2023). *flextable : Functions for Tabular Reporting*. https://CRAN.R-project.org/package=flextable

Gümüş, G., & Balcı, F. (2023). Working memory for time intervals : Another manifestation of the central tendency effect. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02324-z

Haman, J., & Avery, M. (2020). *ciTools : Confidence or prediction intervals, quantiles, and probabilities for statistical models* [Logiciel]. https://github.com/jthaman/ciTools

Hellström, Å. (1985). The time-order error and its relatives : Mirrors of cognitive processes in comparing. *Psychological Bulletin*, *97*(1), 35-61. https://doi.org/10.1037/0033-2909.97.1.35

Hollingworth, H. L. (1910). The Central Tendency of Judgment. *Journal of Philosophy, Psychology & Scientific Methods*, *7*, 461-469. https://doi.org/10.2307/2012819

Hox, J. J. (2010). *Multilevel analysis : Techniques and applications* (2. ed). Routledge, Taylor & Francis.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology. General*, *129*(2), 220-241. https://doi.org/10.1037//0096-3445.129.2.220

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020-1026. https://doi.org/10.1038/nn.2590

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R^2_{GLMM} to random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944-946. https://doi.org/10.1111/2041-210X.12225

Kassambara, A. (2020). *ggpubr : Ggplot2 Based Publication Ready Plots*. https://rpkgs.datanovia.com/ggpubr/

Kenney, J. F., & Keeping, E. S. (1962). Root Mean Square. In *Mathematics of Statistics* (3rd ed, p. 59-60). Princeton.

Koller, M. (2016). robustlmm : An R Package for Robust Estimation of Linear Mixed-Effects Models. *Journal of Statistical Software*, *75*(6), 1-24. https://doi.org/10.18637/jss.v075.i06

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package : Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lüdecke, D., Ben-Shachar, M., Patil, I., & Makowski, D. (2020). Extracting, Computing and Exploring the Parameters of Statistical Models using R. *Journal of Open Source Software*, *5*(53), 2445. https://doi.org/10.21105/joss.02445

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance : An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494-1502. https://doi.org/10.3758/s13428-016-0809-y

Mailund, T. (2019). *Introducing Markdown and Pandoc : Using Markup Language and Document Converter*. Apress. https://doi.org/10.1007/978-1-4842-5149-2

Mirman, D. (2014). *Growth curve analysis and visualization using R* (1st éd.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315373218

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *The Royal Society Publishing*, *14*. http://dx.doi.org/10.1098/rsif.2017.0213

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133-142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Olkkonen, M., McCarthy, P. F., & Allred, S. R. (2014). The central tendency bias in color perception : Effects of internal and external noise. *Journal of Vision*, *14*(11), 5-5. https://doi.org/10.1167/14.11.5

R Core Team. (2022). *R: A language and environment for statistical computing* [Logiciel]. R Foundation for Statistical Computing. https://www.R-project.org/

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed). Stata Press Publication.

Raudenbush, StephenW., & Bryk, A. S. (2002). *Hierarchical Linear Models Applications and Data Analysis Methods* (2e éd.). Sage Publications.

Ryan, L. J. (2011). Temporal context affects duration reproduction. *Journal of Cognitive Psychology*, *23*(1), 157-170. https://doi.org/10.1080/20445911.2011.477812

Saarela, T. P., Niemi, S. M., & Olkkonen, M. (2023). Independent short- and long-term dependencies in perception. *Journal of Vision*, *23*(5), 12. https://doi.org/10.1167/jov.23.5.12

Thériault, R. (2022). *rempsyc : Convenience functions for psychology*. https://rempsyc.remi-theriault.com

Ushey, K., & Wickham, H. (2023). *renv : Project environments*. https://CRAN.R-project.org/package=renv

Wedell, D. H., Hayes, W. M., & Kim, J. (2020). Context effects on reproduced magnitudes from short-term and long-term memory. *Attention, Perception, & Psychophysics*, *82*(4), 1710-1726. https://doi.org/10.3758/s13414-019-01932-z

Wickham, H. (2016). *ggplot2 : Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr : A grammar of data manipulation* [Logiciel]. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Henry, L. (2023). *purrr : Functional programming tools* [Logiciel]. https://CRAN.R-project.org/package=purrr

Xie, Y. (2023). *knitr : A general-purpose package for dynamic report generation in R* [Logiciel]. https://yihui.org/knitr/

Xie, Y., Dervieux, C., & Riederer, E. (2020). *R markdown cookbook*. Chapman and Hall/CRC. https://bookdown.org/yihui/rmarkdown-cookbook