Article:

Wang, Yaqi, Gaskell, Gareth orcid.org/0000-0001-8325-1427 and Gennari, Silvia Patricia orcid.org/0000-0002-2242-4002 (2024) Influences of learned verbal labels and sleep on temporal event memory. Journal of Memory and Language. 104529. ISSN: 0749-596X

https://doi.org/10.1016/j.jml.2024.104529

Journal of Memory and Language 138 (2024) 104529

Contents lists available at ScienceDirect

Journal of Memory and Language

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

Influences of learned verbal labels and sleep on temporal event memory

Yaqi Wang[a,c], M. Gareth Gaskell[b], Silvia P. Gennari[b,*]

[a] School of Foreign Languages and Literature, Shandong University, China
[b] Department of Psychology, University of York, United Kingdom
[c] Centre for Clinical Neurolinguistics, Shandong University, China

A B S T R A C T

Conceptual knowledge is known to modulate episodic memory, but it remains unclear whether and how verbal labels shape event learning and recollection over time. To investigate this issue, we asked participants to study and memorise unfamiliar animations and their titles. The titles conveyed fast or slow motion speed (e.g., *a bus* vs *ambulance travelling*). Event memory was assessed at different time points—soon after learning and after 12 h of sleep or wakefulness—using a timed mental event reproduction task and verbal recall. Unlike previous findings with these stimuli, we found that intentional title study elicited title-related biases on reproduced durations soon after learning. Post-sleep but not post-wakefulness recollection also showed title-related biases and systematically longer reproduced durations. Nevertheless, reproduced durations correlated with stimulus segments, stimulus durations and verbal recall, indicating that event memories combined episodic and verbal conceptual features. Results suggest that intentional verbal learning promoted conceptual influences at encoding and that sleep-dependent consolidation enhanced these influences. We argue that the degree of integration between conceptual and episodic features determines the extent of conceptual influences and, more generally, the role of verbal labels in event learning and memory.

Introduction

Event memory representations are thought to vary in their degree of specificity and distinctiveness. Episodic representations, e.g., the recollection of one's recent breakfast, include experience-specific spatial and temporal features distinguishing them from other similar events. In contrast, representations in semantic memory, e.g., *having breakfast*, convey typical features and action sequences abstracted across multiple experiences, often referred to as *schemas* (Shank & Abelson, 1977). A central goal in the study of human memory is to establish how these semantic and episodic representations relate to each other (Tulving, 1984). For example, much research has been devoted to examining how schemas are formed over time or how they may modulate the encoding and recollection of new experiences (Alba & Hasher, 1983; Gilboa & Marlatte, 2017; van Kesteren et al., 2012). Here, we use verbal labels to introduce contrasting event schemas and examine how they modulate the recollection of dynamic visual stimuli before and after sleep.

Language often accompanies visual perception and learning in daily life, e.g., when learning new information at school, but little progress has been made in understanding the mechanisms underpinning the interactions between verbal and visual event stimuli. Most previous research on the role of verbal labels in visual encoding has typically focused on object or colour categories and static scenes (Carmichael et al., 1932; Feist & Gentner, 2007; Lupyan, 2008; Regier & Kay, 2009; Yuan et al., 2024). One study, for example, compared object labelling with pleasantness judgments and found that recognition memory was poorer after verbal categorisation (Lupyan, 2008). Poorer memory performance was also found in another study comparing the recognition of ambiguous pictures presented with or without spatial sentences: more false alarms were observed when language accompanied the pictures (Feist & Gentner, 2007). These studies argued for an interactive encoding account according to which labelling or the presence of language at perception distorts representations towards the features highlighted by language, resulting in impaired memory performance. Other studies, in contrast, have shown that object labelling may lead to improved object recognition in some cases and argue for more nuanced explanations depending on task demands and stimulus characteristics (Richler et al., 2011, 2013).

Comparatively, fewer studies have examined the role of verbal labels in event memory. Unlike object perception, event processing unfolds over time and relies on dynamic structuring mechanisms guided by top-down event schemas and bottom-up stimulus changes (Zacks, 2020; Zacks et al., 2007). This temporal organisation shapes the nature and quantity of the information subsequently recalled (Hanson & Hirst, 1989; Radvansky & Zacks, 2017). For example, the number of segments or units identified at perception correlates with the information verbally recalled (Hanson & Hirst, 1989; Newtson & Engquist, 1976; Zacks, 2020). These units are organised around the agent's goals and changes

* Corresponding author.
E-mail address: Silvia.gennari@york.ac.uk (S.P. Gennari).

https://doi.org/10.1016/j.jml.2024.104529
Received 7 September 2023; Received in revised form 4 April 2024; Accepted 10 April 2024
Available online 26 April 2024

in spatiotemporal properties such as travel paths and locations (Lichtenstein & Brewer, 1980; Zacks et al., 2009).

In such dynamic contexts, accompanying verbal labels, such as video titles, may offer semantic information supporting event structuring and understanding (Huff & Schwan, 2008; Sakarias & Flecken, 2019). Huff & Schwan (2008), for example, presented unfamiliar videos (coloured billiard balls racing) on their own or preceded by descriptive statements (e.g., *The red ball starts behind, overtakes other balls and wins*). They found that verbal information helped recognition memory discrimination. It was argued that the language helped stimulus structuring and provided an easy-to-access schema (Huff & Schwan, 2008). This finding confirmed previous text-based studies in which a title conveying an event schema facilitated the comprehension and recollection of unfamiliar texts compared to an untitled condition (Bransford & Johnson, 1972).

Verbal labels accompanying dynamic stimuli may also bias encoding towards conceptual features, as argued by the interactive encoding account, leading to less accurate memory (Lupyan, 2008; Feist & Gentner, 2007). This possibility is consistent with numerous studies demonstrating influences of prior knowledge on memory for continuous stimulus dimensions such as size, spatial location, or duration, for which memory traces are typically inexact (Bonasia, Blommesteyn, & Moscovitch, 2016; Hemmer & Steyvers, 2009a,b; Huttenlocher, Hedges, & Bradburn, 1990; Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, & Vevea, 2000; Tompary & Thompson-Schill, 2021). For example, recalling the size of previously studied objects such as fruits and vegetables is modulated by prior category knowledge, biasing recalled sizes towards typical category values (Hemmer & Steyvers, 2009b). Although less studied, temporal event memory is also generally inexact relative to stimulus durations because recalling the time course of events involves reconstructing the encoded units and details rather than judging stimulus duration as such (Arnold et al., 2016; Faber & Gennari, 2015; Jeunehomme et al., 2017; Jeunehomme & D'Argembeau, 2020; Wang & Gennari, 2019). For example, verbally recalling one's trip from one campus building to another takes less time than the actual trip. Instead, it involves mentally re-enacting the path and landmarks of our travel in a spatially and temporally compressed fashion (Jeunehomme et al., 2017; Bonasia et al., 2016; Arnold et al., 2016). Thus, recalling the time course of videos showing objects travelling, as other inexact domains, might be biased by the knowledge of what those objects are and how fast they typically move, e.g., an ambulance or a bus.

However, little research has examined whether and how labels at encoding may modulate visual event learning and subsequent dynamic recollection. Two previous studies examining whether language-induced event conceptualisations modulate visually cued event recollection have failed to detect temporal biases (Wang & Gennari, 2019). These studies asked viewers to read and understand videos titles implying different motion speeds—e.g., *a firework rocket launched into the sky* vs *a Chinese lantern rising up into the sky.* These titles preceded ambiguous geometric animations, e.g., a square rising, as shown in Fig. 1, and thus, provided relevant event schemas to conceptualise and understand the unfamiliar stimuli. This manipulation was expected to shape the animation's encoding, leading to title-induced biases in temporal recollection, i.e., relatively longer or shorter recollection as a function of title. Contrary to expectations, the results suggested that title understanding at encoding did not modulate the stored event representations.

The failure to detect encoding interactions between verbal and visual stimuli in Wang & Gennari (2019) may be due to several reasons. One possibility is that conceptualising the videos in title terms led to relatively shallow processing of the titles' meanings or weak semantic associations between the titles and the videos. For example, participants may have understood that the square in Fig. 1 was a rocket or a lantern but did not grasp the full motion speed implications of this concept or did not apply them to the moving figure. Another possibility is that verbally induced biases emerge after delays, e.g., recalling the animations a day after encoding, because other memory processes intervene that may transform the association between titles and videos (Dudai, 2012). Thus, more research is required to understand whether and how verbal labels interact with visual stimuli at encoding.

The present study

Here, we investigated the role of explicitly learned verbal labels in immediate and delayed temporal recollection. We specifically examined whether joint learning of verbal and visual stimuli promotes temporal biases soon after learning or after a delay with and without intervening sleep. Using the stimuli in Fig. 1 from Wang and Gennari (2019), the present study specifically encouraged the encoding of video titles together with the animations for upcoming memory tests targeting both animations and phrases. The task instructions thus differed from those in Wang and Gennari (2019) in emphasising title learning or memorisation rather than language-based event understanding. It is well-known that the nature of the operations performed during learning (deep vs shallow processing) and study intent (e.g., intentional study vs incidental learning) modulate the strength of memory traces (Craik & Tulving, 1975; Oberauer & Greve, 2022; Popov & Dames, 2022; Silverstein & Marshall, 1968). These two concepts are interrelated because participants' intention to remember the paired stimuli promotes deeper encoding than incidental stimulus processing, e.g., by establishing more elaborate semantic links for the stimuli, which facilitates subsequent recollection (Einstein & Hunt, 1980; Staresina, Gray, & Davachi, 2009; Anderson & Reder, 1979). Based on these findings, we hypothesised that explicitly instructing participants to learn both titles and animations would promote deeper encoding and, thus, the processing of more semantic links between them than those in previous results. These additional links would bolster the semantic integration between titles and videos during encoding, thus increasing the likelihood of temporal biases at recollection.

Sleep-dependent consolidation can also increase the semantic integration of titles and animations, as memory consolidation is known to strengthen and stabilize newly formed memories over time (Dudai, 2012). Memory consolidation is a brain process integrating newly

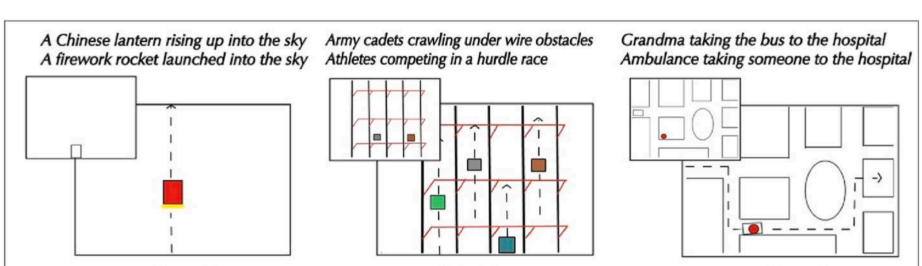

Fig. 1. Example stimuli. Only one of the alternative titles was presented to a participant. Smaller rectangles represent the frame near the beginning used to cue recollection—the *cue frame*. Larger rectangles represent the animations. The arrows indicate motion path but were not shown to participants.

acquired memories in subcortical regions with cortical structures storing semantic knowledge, and sleep is thought to promote this process (Dudai, 2012; Klinzing et al., 2019; Landmann et al., 2014; Rasch & Born, 2013; Stickgold & Walker, 2013). Behaviourally, sleep results in better recall and recognition memory performance than an equivalent wakefulness period (Berres & Erdfelder, 2021). For example, more event details and paired associated stimuli are recalled after sleep than after wakefulness (Aly & Moscovitch, 2010; Ekstrand, 1977; Feld & Born, 2017; Lau et al., 2010; Rasch & Born, 2013; Wolford, 1971). Sleep has also been shown to benefit schema-related information compared to schema-unrelated information (Hennies et al., 2016) and promote the integration of newly learned stimuli with prior knowledge (Dumay & Gaskell, 2007). In comparisons across sleep and wakefulness periods, however, memory differences may not only stem from sleep-dependent consolidation but also from other processes taking place during wakefulness: Wakeful activities can interfere with the retrieval and consolidation of previously learned stimuli, thus impairing subsequent recollection of unconsolidated memories (Dewar et al., 2012; Wixted, 2004, 2005; Yonelinas et al., 2019).

Based on this prior research, our study investigated whether explicit instructions to learn video titles modulate visual event memory soon after learning or after delay periods with or without intervening sleep. Specifically, we ask whether title-related biases are observed after sleep compared to wakefulness and whether additional changes emerge compared to recollection soon after learning. Sleep-dependent consolidation should strengthen the title-video association and semantic integration relative to wakefulness, leading to title-related biases and better stimulus recollection. In contrast, the interference of wakeful activities may induce forgetting during wakefulness, resulting in weaker title biases and impaired recollection. However, if joint learning of titles and animations promotes semantic integration and title-related biases soon after learning, as hypothesised above, sleep-dependent consolidation may further modulate the already integrated verbal and visual stimuli. Therefore, other modulations may emerge after sleep compared to recollection soon after learning.

To address these hypotheses, four participant groups were repeatedly exposed to animations and accompanying titles using a similar paradigm as Wang and Gennari (2019) (see Fig. 1). Participants were instructed to study *both* the titles and the animations for later memory tests of the animations and phrases. Unlike timing studies, no instructions to attend to time or stimulus duration were presented at any point during the study (Grondin, 2010; Matthews & Meck, 2016). The memory tests included a task targeting temporal memory and a verbal recall task. Temporal memory was assessed with visually cued episodic reproductions or mental replays—replaying a previously experienced event in the mind's eye (D'Argembeau et al., 2021; Faber & Gennari, 2015; Schacter & Addis, 2007, 2009). The duration of these mental reproductions has been shown to correlate with event structure and verbal recall, suggesting that temporal memory is grounded on the encoded event representations (Faber & Gennari, 2015; Jeunehomme & D'Argembeau, 2020; Jeunehomme et al., 2017). For example, the number of segments or units in the stimuli correlates with the reproduced duration: the more units are identified, the longer the mental reproductions (Faber & Gennari, 2015; Jeunehomme et al., 2017). Moreover, the more words are used in recalling what happened in an event, the longer the reproduced duration (Wang & Gennari, 2019). Therefore, despite often being inaccurate relative to stimulus duration, reproduced durations reflect properties of the encoded event.

In our study, episodic reproductions and verbal recall occurred at different times after learning, depending on the group (see Fig. 2). The Sleep group learned the stimuli in the evening and was tested in the morning of the following day, whereas the Wake group learned the stimuli in the morning and was tested in the evening of the same day. Two other participant groups were tested 10–12 min after learning within the same experimental session. This session occurred in the morning or the evening to control for possible time-of-day effects on the Sleep and Wake groups' performance (Folkard, 1979; Tilley & Warren, 1983). These groups are specified as the *Morning* and *Evening* groups in Fig. 2 and collectively referred to as the *Immediate* groups, in contrast to *Delay* groups.

Based on our hypotheses, we expected that the different study instructions relative to previous studies might lead to deeper learning of title-video pairs and, thus, elicit title-induced temporal biases in immediate groups, i.e., shorter or longer episodic reproductions as a function of title. Moreover, title-induced temporal biases, better stimulus recollection, and possibly other influences may emerge after sleep compared to wakefulness and immediate recollection since sleep should promote more robust and deeper title-video integration. To gain insights into the memory representations entertained after delays, we also examined whether delayed episodic reproductions correlate with verbal recall, stimulus segments and stimulus duration in similar ways as in immediate recollection. Finally, if the change of study instructions relative to Wang and Gennari (2019)'s studies is responsible for deeper semantic integration and title-related biases, corresponding differences should be observed when comparing study instructions. We explored this possibility by comparing previous results in Wang and Gennari (2019) to the present Morning and Evening groups.

Methods

Data availability

All stimuli, data and results associated with the present manuscript can be found on Open Science Framework (OSF) at https://osf.

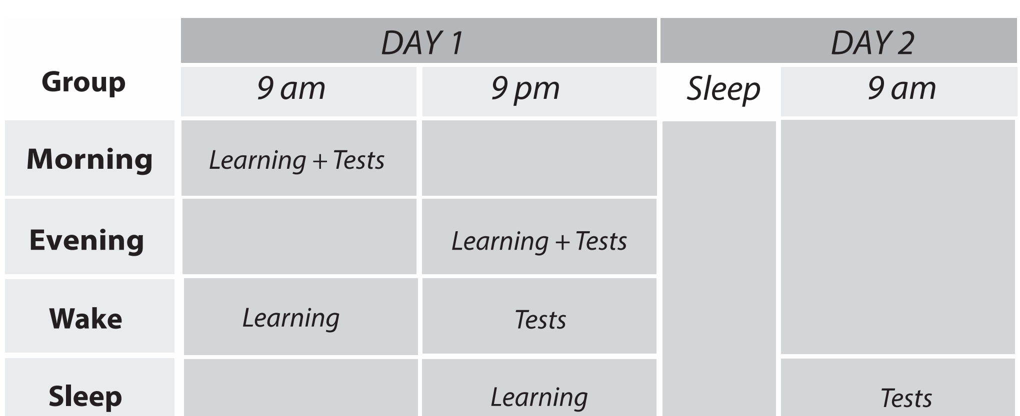

Fig. 2. Learning and testing sessions in each group. The testing session included the *mental replay* and *verbal recall* tasks in this order.

io/k8yxa/?view_only. This source contains verbatim learning instructions (including those in Wang & Gennari, 2019), Supplementary Methods, Supplementary Results and the R code producing the statistical results.

Participants and groups

A total of 264 English monolinguals allocated to four participant groups were recruited from the University of York and tested in lab conditions. Based on prior sleep and duration memory studies, we aimed to recruit 60–65 participants per group (Dumay & Gaskell, 2007; Wang & Gennari, 2019). The Wake and Sleep groups were tested first, with participants randomly allocated to either group. The Morning and the Evening groups were tested later. Participants in these groups were also randomly assigned to either group after agreeing to participate in the study. In total, 14 participants were excluded because their responses were inaccurate in more than 30 % of recall trials or did not follow instructions (e.g., they pressed the key to skip the replay task). The verbal recall data of three participants was lost due to software failure, so only reproduced duration was analyzed for these participants. After these exclusions, the participant numbers were as follows: Sleep group: N = 62; Wake group: N = 64; Morning group: N = 63; Evening Group: N = 61.

Stimuli

The stimuli (21 animations and 42 phrases) were taken from Wang and Gennari (2019). The animation stimuli systematically varied in duration from 3 to 9sec in increments of a second. There were three animations for each of the seven possible durations. The titles differed in the nouns used to refer to the moving objects (*rocket* vs *lantern*), except for seven animations for which the verb conveyed the speed (*stroll* vs *run)*. A series of online questionnaires were conducted to control for confounding stimulus characteristics. These online questionnaires checked that (1) the phrases indeed implied contrasting motion speed (motion speed rating), (2) the two speed-related phrases (e.g., *firework rocket* vs *lantern*) applied to the animation equally well (phrase-animation fit rating), (3) the familiarity of the events referred to by the phrases did not differ in familiarity (familiarity ratings), and (4) the perception of the scene scale (being closer or further away) did not change across conditions. These questionnaires had the same design as the primary studies below. Table 1 reports the mean and SD of these ratings. Except for the motion speed questionnaire, all other comparisons were not significant (see Supplementary Methods, in OSF). Additional segmentation studies reported by Wang and Gennari (2019) also indicated that titles did not alter the number of segments perceived in the stimuli.

Design and procedure

Each of the 21 animations was paired with two possible titles (fast vs slow). Each title-animation pair was assigned to a different stimulus list, resulting in two stimulus lists. Each list contained approximately half of the animations in either the slow- or fast-title condition. Each participant thus saw an animation in one title condition but saw all animations across title conditions (see data files). All groups performed identical tasks.

Table 1
Mean and standard deviations from pre-test rating studies.

Rating Tests	Condition	
	Slow	Fast
Phrase implied speed	3.17(.92)	5.12(.92)
Phrase-animation fit	6.17(.39)	6.29(.40)
Event familiarity	4.83(1.42)	5.33(1.00)
Perceived scene scale	3.74(1.07)	3.92(1.13)

Learning task. Immediately before the learning session, the instructions encouraged participants to study the verbal and visual stimuli for later memory tests targeting both types of information (see *Instructions* file in OSF). During the task, participants saw all animations and corresponding title phrases in random order three times, with an intervening screen between cycles. The trial structure is depicted in Fig. 3. For each of the 21 title-animation pairs in a list, participants first saw the titles at the top with the first animation frame underneath (the cue frame). When they pressed a key, the animation was shown in full. After participants saw all stimuli, another viewing cycle was announced to promote better learning.

Memory tasks. A mental replay and a verbal recall task were included. In the replay task, participants were asked to reproduce the animations in their minds *exactly as they occurred in their original time course* when prompted with the cue frame. They were asked to click the mouse at the beginning and end of their replay (see Fig. 3). The reproduced duration was measured from the START to the STOP click. Note that only a visual image (the cue frame) was presented at retrieval to cue an animation's recollection. This feature sets our study apart from others investigating retrieval biases elicited by verbal information at retrieval (Loftus, 2005; Loftus & Palmer, 1974; von Sobbe et al., 2021; Wang & Gennari, 2019). The verbal recall task, which was announced as the last task of the study, requested "as many details as they could remember about the animations (including physical characteristics, e.g., shape movements and colour changes)". Participants were instructed not to guess, as the study was interested in what they did or did not remember. This reminder allowed us to exclude from the duration reproduction data those recall trials indicated as forgotten or left blank. Recollection of the titles was not explicitly requested, although participants were initially told that title-phrase recollection would be tested. Verbal responses were typed into a screen box (see Fig. 3).

Group testing and sleep questionnaires. Participant groups differed in the time of learning and testing, as shown in Fig. 2. Participants in the Morning and Evening groups performed the learning and memory tasks in the same session with an intervening distractor task (10–15 min of maths questions). This session lasted approximately 1 h. Participants in the Wake and Sleep groups performed the learning task in the morning and evening, respectively, and were tested 12 hrs later in a separate lab session. Participants in these groups agreed to abstain from alcohol, drugs and caffeine and to follow their regular sleep hours or activities during the intervening delay. They were reminded of this commitment after the learning session.

All groups filled out three *questionnaires* that provided measures of sleep patterns and alertness (Pittsburgh Sleep Quality Index, Morningness-Eveningness Questionnaire, Stanford Sleepiness Scale

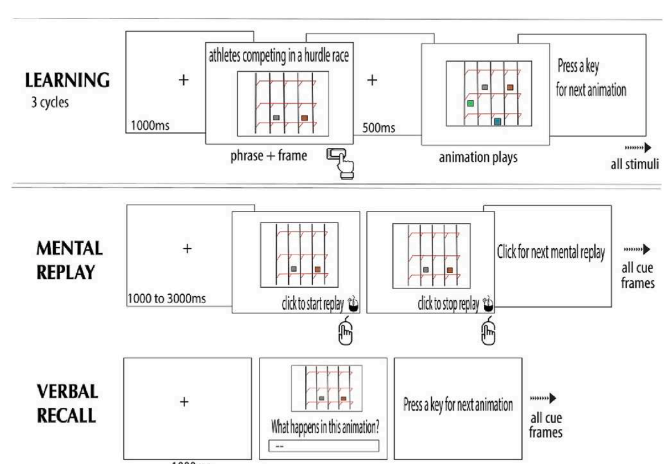

Fig. 3. Trial structure in all tasks. The reproduced duration was computed from the start and stop click of the mental replay task.

(Buysse et al., 1989; Hoddes et al., 1973; Horne & Ostberg, 1976). The Sleep Quality Index and Morningness-Eveningness questionnaires were filled out after testing, whereas the sleepiness scale was indicated before testing. The mean and standard deviations for each questionnaire are shown in Table 2. Anovas with groups as factor indicated non-significant differences (Pittsburgh Sleep Quality Index: $F(3, 243) = .56$, p=.64; Morningness-Eveningness Score: $F(3, 243) = .99$, p=.40; Stanford Sleepiness rating: $F(3, 243) = 2.17$, p=.09). Importantly, comparisons across the Sleep and Wake groups were not significant (Pittsburgh Sleep Quality Index: $t(124) = 1.38$, p=.17; Morningness-Eveningness Score: $t(124) = .47$, $p = .64$; Stanford Sleepiness rating: $t(124) = 1.04$, $p = .30$). These questionnaires indicated similar levels of alertness across groups, and no significant differences in chronotypes or sleep quality (e.g., keeping regular hours, number of hours sleeping). None of the participants reported regularly napping. Nevertheless, we did not specifically ask participants to report what they did during the wakeful period.

Data treatment and analyses

Reproduced duration. We excluded trials faster than 1 s, which were likely errors, and those cases indicated as forgotten or left blank in the verbal recall task. To minimize the influence of outliers in all our measures, we used the Tukey method for outlier removal for each group or individual distribution, i.e., values larger than $Q3 + (1.5 \times IQR)$ were trimmed, where $Q3$ is the third quartile in the data, and IQR is the interquartile range (quartile 3 — quartile 1) (Tukey, 1977). Linear mixed-effects models were computed in the *R's lme4* and *lmerTest* packages (Bates et al., 2015; Kuznetsova et al., 2017). Models predicting reproduced durations included stimulus duration and the number of stimulus segments as fixed factors. The number of segments in each animation was obtained by a previous segmentation task, as reported in Wang & Gennari (2019). These variables have been shown to account for variance in memory-based reproduced durations (Faber & Gennari, 2015; Jeunehomme et al., 2017). The models also included by-participants, by-items random factors, title condition, and stimulus duration slopes when convergence allowed (Barr et al., 2013). To simplify the models and test our main hypotheses, we first compared the Morning and Evening groups to examine possible time-of-day effects. If there were no relevant differences, we treated these groups as the *Immediate group*. The main models used Helmert contrasts for the three groups (Immediate, Wake, Sleep) comparing Immediate vs Delayed testing (coded as −2 1 1) and Wake vs Sleep groups (coded as 0—1 1). Slow and Fast title conditions were coded 1 and −1. Post hoc contrasts used *R*'s *emmeans* package.

Verbal recall. Since our hypotheses primarily concern the effect of video titles on reproduced duration, we did not attempt to manually code event recollection to assess how the recollection quality differs across delay conditions. Much memory research has shown differences across various delay periods (Wixted, 2004, 2005; Berres & Erdfelder, 2021; Aly & Mostcovitch, 2010, Sekeres et al., 2016). Instead, we coded title recall and computed the number of words recalled as a proxy for the overall amount of information remembered, as in Wang & Gennari (2019). This measure was then correlated with reproduced duration to establish whether recalled information remains associated with reproduced duration after delays. Recall responses were considered inaccurate and excluded from analyses when no correct information was provided about a stimulus. Due to the extensive learning task, all groups had an average accuracy of 99 %, with participants' accuracy ranging from 86 % to 100 %. In a few cases, participants reported a single feature, e.g., *rocket* or *lantern*, which was considered accurate.

Table 2
Mean scores across groups for sleep-related questionnaires.

Questionnaire	Morning	Evening	Wake	Sleep
Pittsburgh Sleep Quality Index	6.7(3.2)	6.4 (2.9)	6 (2.7)	6.3 (3.1)
Stanford Sleepiness Scale	2.5 (0.9)	2.7 (1)	2.8 (1.3)	3 (1.1)
Morningness-Eveningness Questionnaire	47.8 (9)	47.5 (9.6)	47.8 (9.3)	50 (8.4)

Because there was considerable variability across groups in the number of words and title reports, we aimed to reduce it in two ways. First, we identified outliers in each group as indicated above and replaced them with the corresponding group median (extreme values in this data are not considered performance errors, so replacement was appropriate). This treatment proved more appropriate than log values because the model assumptions were better met (e.g., the model residuals were nearly normally distributed), although both measures produced equivalent results. Second, we computed the number of unique meaningful words (besides total word counts) by automatically removing function words and ignoring word repetitions. This measure minimised individual and group differences in wordiness and title recall. Note that it was not always possible to separate the title report from other stimulus information, and therefore, group differences in title reports would result in word count differences. By ignoring word repetitions in a trial, repeated references to the same moving entity—for example, in the title and the remaining verbal recall—could be counted once. Finally, counting unique meaningful words approximates the number of event details reported: Unique meaningful words such as nouns, verbs, adjectives and adverbs refer to actors, actions and their properties (e.g., colours, moving direction).

Unique, meaningful word counts. To compute this measure, we proceeded in three steps: For each verbal report, we first manually removed introductory phrases (e.g., *the title of this* animation *was* …) or parenthetical comments directed to the experimenter (e.g., *I do not remember any other colour change)*. Second, we used scripting in Excel to remove a list of function words from each text response (e.g., *the, a, some, it, of, that, which, this*, *these*, *be, been, is, are, has, have*, etc.). Finally, additional scripts removed all duplicate words. The linked data files contained both unedited and modified trials. Note that these word counts could include title words if a title was reported in a trial (see Table 3). In many trials, it was not possible to separate the title from title-unrelated information because the title was the only information reported. Nevertheless, if the title contained a reference that was later repeated in the description, this reference was counted only once.

Title Reporting. To ascertain whether there were differences across groups in title reporting, we manually coded whether each recall trial started with or only contained the title. Recall that although learning instructions anticipated the testing of verbal information, participants were not explicitly instructed to report the titles in the verbal recall task. Thus, mentioning the title at the trial start suggests that the title was highly available upon seeing the visual cue. This possibility is demonstrated by paired-associates learning studies where a pair member makes the other member highly available after learning (Tulving & Pearlstone, 1966) and is consistent with multiple language production studies showing that more accessible elements in memory are produced first (Bock, 1982, 1987a, 1987b; Bock & Warren, 1985; McDonald et al., 1993).

In coding the verbal recall data, we consider that the title headed a recall trial if (a) the main syntactic components of the title (the noun, the verb and adverbial phrase) or near synonyms were mentioned in the title order and (b) the phrase was diagnostic of the title seen at encoding, i.e., the critical speed information was indeed implied (see examples Table 3). This strategy excluded references that could equally apply to the fast and the slow title conditions. For example, descriptions such as *something flying up, people exercising,* or *someone traveling to the hospital* for the animations in Fig. 1 would not be counted as title recollection. In most cases, participants used commas or periods to separate the title from other details or simply provided the title as a description of what happened in the animation (see data files).

Table 3
Examples of typed recall for the three animations in Fig. 1.

Title	Verbal recall examples	Code
A Chinese lantern raising up into the sky	1. A Chinese lantern going up into the sky, it becomes yellow and then orange as it rises, and the bottom is red.	1
	2. The lantern glows red and yellow as it is lit, and floats into the air in a wiggly line.	0
	3. The lantern flashes yellow at the bottom, turns red, then moves to the top of the screen.	0
Army cadets crawling under wire obstacles	1. Army cadets completing an obstacle course, brown and grey square move fairly steadily through the course and then a fast blue square catches up. Another square follows the other three, but the brown square finishes first	1
	2. The cadets move upwards under the wire. The cadets appear shaky as they are doing so. More cadets come onto the screen and also move upwards shakily. As the cadets go under the wire, the wire changes to a grey colour	0
	3. The grey square finishes first, followed by brown. A lime green square appears, and a darker green square also appears, which finishes last	0
An ambulance taking someone to the hospital	1. An ambulance taking someone to a hospital, the rectangle on the left picks up the red dot and takes it to the middle rectangle on the right	1
	2. The ambulance turns right, then left. Picks up the red circle. Continues straight, turns left, and goes around the oval from the top. Drops the red circle in the second square on the right. The box changes to orange.	0
	3. The ambulance (white square) moves forward, the right and then left and stops at the patient's house. The red circle is the patient, who moves from its house to the ambulance. The ambulance then carries the patient forward and turns left at the end of the patient's house, in-between the house and the oval. At the top of the oval, the ambulance turns right and then immediately right, going between the hospital and the oval. The ambulance stops at the hospital, and the patient moves from the ambulance to the hospital.	0

Table 4
Reproduced Duration Results (sec).

1. Evening & Morning Groups			
Predictors	Estimates	CI	p
(Intercept)	4.61	4.02 – 5.21	**<0.001**
Stimulus Duration	0.68	0.49 – 0.86	**<0.001**
Stimulus Segments	0.68	0.49 – 0.86	**<0.001**
Title (Slow v Fast)	0.18	0.02 – 0.33	**0.029**
Group (Evening v Morning)	−0.02	−0.39 – 0.35	0.911
Title * Group	0	−0.20 – 0.20	0.981
Random Effects			
σ^2	1.63		
$\tau_{00\ \mathrm{SubjID}}$	1		
$\tau_{00\ \mathrm{itemN}}$	0.16		
$\tau_{11\ \mathrm{SubjID.duration}}$	0.03		
$\tau_{11\ \mathrm{itemN.Title}}$	0.02		
$\rho_{01\ \mathrm{SubjID}}$	−0.45		
$\rho_{01\ \mathrm{itemN}}$	−0.49		
ICC	0.41		
N_{SubjID}	124		
N_{itemN}	21		
Observations	2471		
Marginal R^2 / Conditional R^2	0.201 / 0.531		
2. Immediate, Sleep & Wake Groups			
Predictors	**Estimates**	***CI***	***p***
(Intercept)	3.22	2.61 – 3.82	**<0.001**
Stimulus Duration	0.3	0.21 – 0.38	**<0.001**
Stimulus Segments	0.2	0.09 – 0.31	**<0.001**
Title (Slow v Fast)	0.06	0.02 – 0.10	**0.005**
Group1 (Immediate v Delay)	0.13	0.04 – 0.23	**0.004**
Group2 (Wake v Sleep)	0.34	0.15 – 0.54	**0.001**
Title * Group1	−0.02	−0.04 – 0.01	0.239
Title * Group2	0.06	0.00 – 0.11	**0.037**
Random Effects			
σ^2	1.84		
$\tau_{00\ \mathrm{SubjID}}$	1.34		
$\tau_{00\ \mathrm{itemN}}$	0.12		
$\tau_{11\ \mathrm{SubjID.duration}}$	0.03		
$\rho_{01\ \mathrm{SubjID}}$	−0.47		
ICC	0.46		
N_{SubjID}	250		
N_{itemN}	21		
Observations	4954		
Marginal R^2 / Conditional R^2	0.158 / 0.546		

Results

Reproduced duration from memory

We hypothesized that the study instructions would lead to deeper learning of title- video pairs and elicit title-induced biases in reproduced duration. To test this hypothesis in the Immediate groups and furthermore check whether the time of day modulates reproduced duration, we first compared the Morning and Evening groups. As shown in Table 4, the Evening and the Morning groups did not differ significantly, but both groups showed a title effect. Animations learned with titles implying fast motion were reproduced shorter overall than those implying slow motion (see Fig. 4). More complex models examining possible interactions with stimulus duration indicated no significant interactions, suggesting that Morning and Evening groups did not differ across stimulus durations. Based on these results, we combined the Evening and the Morning groups into the Immediate group in subsequent models.

We also hypothesized that title-induced biases might be more robust after sleep than after wakefulness. Our results in Table 4.2 indicated an overall effect of Title and a Title by Group interaction, indicating that although there was an overall Title bias in reproduced durations, the Title effect was absent in the Wake group. This result was confirmed with post hoc Title contrasts using *emmeans* package (Immediate Group t = -2.7, p = .009; Wake Group t = -2.3, p = .7; Sleep Group t = -2.7, p = .008). The mean difference between the Slow and Fast conditions was 182 ms for the Immediate groups and 197 ms for the Sleep group. See Fig. 4B.

The overall pattern of reproduced duration in Fig. 4A replicates previous findings with this paradigm showing inaccurate or distorted reproductions relative to stimulus durations: shorter animations in the stimulus set are lengthened, whereas longer animations are shortened (Roy & Christenfeld, 2008; Wang & Gennari, 2019). For example, three-second stimuli are reproduced as longer, whereas nine-second stimuli are reproduced as shorter. This pattern emerges from the systematic variability in the stimulus durations making shorter videos appear longer in the set (and vice-versa) (see also Huttenlocher, Hedges, & Duncan, 1991; Huttenlocher, Hedges, & Vevea, 2000; Roy & Christenfeld, 2008; Wang & Gennari, 2019 for further discussion). Interestingly, the Sleep group shows the same distortions as the Immediate group shifted upwards, on average 660 ms longer, as indicated by the main effects of group in Table 4.2. Reproduced durations after sleep were also 740 ms longer than after wakefulness, and Delay reproductions were longer than Immediate ones (see Fig. 4). The mean reproduced durations were 560 ms, 552 ms, and 626 ms for the Immediate, Wake, and Sleep groups, respectively.

There were also differences between the Immediate and Wake groups in the size of the distortions. As shown in Fig. 4A, the Immediate and Wake groups display opposite trends for shorter vs longer stimuli. This pattern was revealed by additional models, including interactions with stimulus duration (see SR-Table 1 in Supplementary Results). The Wake group exacerbated the distortions typically observed in these paradigms, i.e., under-reproduction of longer stimuli and over-reproduction of

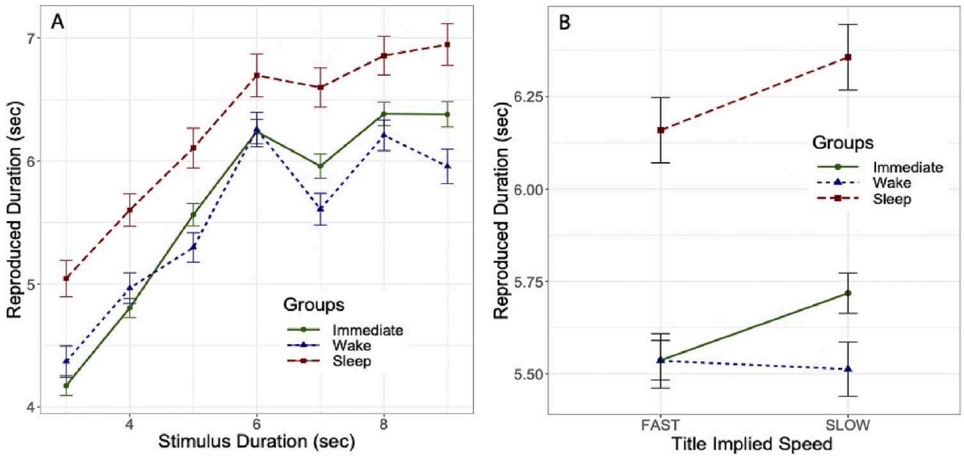

Fig. 4. Panel A: Reproduced Duration by group and stimulus duration. Panel B: Mean Reproduced Duration by group and Title condition. Error bars show standard errors.

shorter ones, suggesting less accurate reproductions in the Wake group relative to the Immediate group.

These results confirm our expectations that Title-induced biases would occur after intentionally studying the titles with the animations. Likewise, they support the expectation that title-induced temporal biases would occur after sleep compared to wakefulness and that additional modulations might be observed relative to immediate recollection. The specific nature of post-sleep influences relative to the Immediate groups—the lengthening of reproduced durations—was not anticipated. We revisit this issue in the general discussion.

Verbal recall and relationship to reproduced duration

Based on prior sleep research indicating differential memory performance after sleep, one might expect differences in the amount of information recalled, as measured by the number of unique meaningful words recalled. Moreover, we ask whether reproduced durations correlate with this measure when a wakeful delay or sleep intervenes, as previously shown for reproductions soon after learning. To examine these predictions, we ran two statistical models. The first model compared unique, meaningful word counts across groups, and the second one correlated these counts with reproduced durations. In these models, we collapsed across Title conditions (fast vs slow) because our hypotheses only concern broad patterns within and across groups (e.g., differences in the information reported and its correlations with recollection). Finally, we expected that title reports would be higher after sleep than after wakefulness, particularly at the beginning of the trial, as this use indicates the strength of their association with the visual cue (Tulvin and Pearlstone, 1966). Generalised linear models were used to compare the probability of reporting the titles across groups to examine this possibility.

Unique Meaningful Word Counts: Table 5 reports descriptive statistics for total and unique meaningful word counts. Table 6 shows the model results comparing recall words across the Immediate, Wake and Sleep groups (using Helmert contrasts). Comparisons across the Morning and Evening groups showed no significant difference and were combined into the Immediate Group (see SR-Table 2 in the OSF Supplementary Results). These tables indicate that the Immediate group used fewer unique meaningful words in their verbal reports than the two Delay groups, whereas the Sleep group used marginally more words than the Wake group. These differences are consistent with the time each group spent on the verbal recall task. The Immediate groups spent an average of 15.5 min on this task, whereas the Wake and Sleep groups spent an average of 19 and 23 min, respectively. We will come back to these differences in the discussion.

Table 5
Descriptive statistics for unique meaningful word and total word counts.

Group	Unique Meaningful Word Count			Total Word Count		
	Mean	SD	Range	Mean	SD	Range
Immediate	12.42	5.85	1–30	22.78	13.44	1–62
Wake	14.30	6.59	1–34	26.90	14.93	1–74
Sleep	15.48	6.43	1–33	30.12	15.43	1–72

Table 6
Comparison of unique content word counts across groups.

Word Count: Immediate, Sleep & Wake Groups			
Predictors	Estimates	CI	p
(Intercept)	−1	−1.53 – -0.47	**<0.001**
Stimulus duration	0.08	0.00 – 0.15	**0.04**
Stimulus segments	0.15	0.05 – 0.24	**0.004**
Groups1 (Immediate vs Delay)	0.13	0.07 – 0.18	**<0.001**
Groups2 (Wake vs Sleep)	0.1	−0.01 – 0.21	0.072
Random Effects			
σ^2	0.41		
$\tau_{00\ SubjID}$	0.38		
$\tau_{00\ itemN}$	0.1		
ICC	0.54		
N_{SubjID}	247		
N_{itemN}	21		
Observations	4895		
Marginal R2 / Conditional R2	0.124 / 0.599		

Relating unique, meaningful word counts and temporal reproductions. A model predicting reproduced durations from unique meaningful word counts, shown in Table 7, indicated a similar overall relationship between these variables across groups. There was no interaction between word counts and groups. Post hoc contrasts indicated that the non-significant difference in the Immediate vs Delay comparison is due to the non-significant difference between the Wake and Immediate groups in this model. As indicated in the previous section, stimulus duration also significantly correlates with reproduced durations. Thus, the relationship between verbal recall, stimulus duration and reproduced duration suggests that despite delays, temporal reproductions do not result from guesses. Instead, they are grounded in event memories and derived from stimulus properties.

Title-phrase reporting. We used generalized mixed effect models to compute the likelihood of reporting the title at the trial start to examine whether there were group differences in title reporting. The model included meaningful word count as a predictor to control for its influence, as using more words may increase the likelihood of reporting the

Table 7
Reproduced duration as a function of unique meaningful word counts and groups.

Reproduction Duration: Word Count and Groups			
Predictors	Estimates	CI	p
(Intercept)	4.11	3.61 – 4.62	<0.001
Stimulus duration	0.28	0.21 – 0.36	<0.001
Stimulus segments	0.24	0.09 – 0.40	0.002
Unique, meaningful word counts	0.21	0.15 – 0.27	<0.001
Group1 (Immediate v Delay)	0.07	−0.02 – 0.17	0.137
Group2 (Wake v Sleep)	0.36	0.16 – 0.57	<0.001
Random Effects			
σ2	1.97		
$\tau_{00\ SubjID}$	1.24		
$\tau_{00\ itemN}$	0.11		
ICC	0.41		
N_{SubjID}	247		
N_{itemN}	21		
Observations	4895		
Marginal R2 / Conditional R2	0.167/0.506		

titles. The initial comparison between the Morning and Evening groups did not differ significantly, so these groups were combined into the Immediate group as before (see SR-Table 3 in the OSF Supplementary Results). The results are shown in Table 8. The Sleep group was more likely to initiate verbal reports with the title than the Wake group, but there was no difference between the Immediate and the Delay groups considered together. Further post hoc contrasts indicated no other significant difference between groups. Numerically, the Wake group was the least likely to report the title phrase (see predicted probabilities in Table 8). These results confirm the hypothesis that compared to the Wake group, the Sleep group had better title recollection than the Wake group.

Together, the results of this section indicate that the Sleep group used marginally more unique meaningful words in their recall and was more likely to report the titles than the Wake group. The Immediate groups used comparatively fewer words than the two Delay groups. The number of unique meaningful words used in each trial was also generally associated with temporal reproductions and predicted longer reproductions for the Sleep than the Wake group. Delay groups considered together did not differ significantly from the Immediate groups in Title reports, suggesting that title recollection in Immediate groups falls in between the Delay groups, as does the relationship between unique meaningful words and reproduced duration (see Fig. 5).

Table 8
Likelihood of reporting Titles at trial start.

Title Phrase			
Predictors	Odds Ratios	CI	p
(Intercept)	1.02	0.60 – 1.71	0.946
Unique, meaningful word count	1	0.88 – 1.14	0.993
Group1 (Immediate vs Delay)	1.06	0.80 – 1.39	0.698
Group2 (Wake vs Sleep)	1.88	1.05 – 3.37	**0.034**
Random Effects			
σ^2	3.29		
$\tau_{00\ SubjID}$	10.21		
$\tau_{00\ itemN}$	0.44		
ICC	0.76		
N_{SubjID}	247		
N_{itemN}	21		
Observations	4895		
Marginal R^2 / Conditional R^2	0.015/0.768		
Groups	**Predicted probabilities**	**95 % CI**	
Immediate	.48	.35 –.61	
Wake	.36	.23 –.52	
Sleep	.67	.48 –.82	

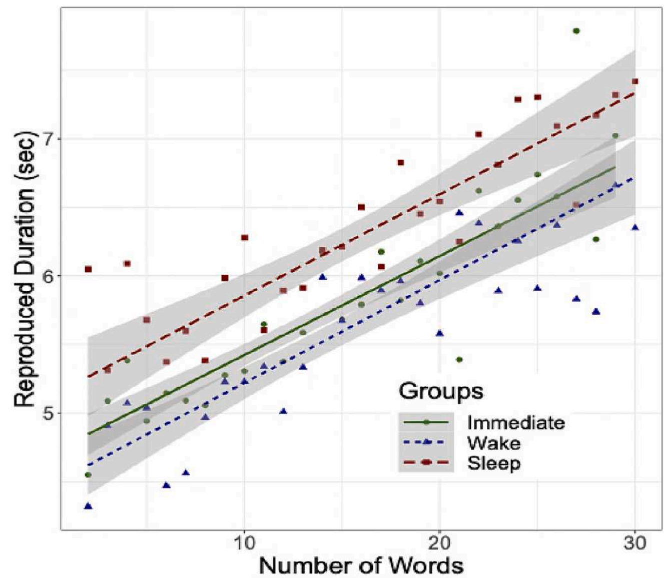

Fig. 5. Relationship between the number of unique meaningful words reported and Reproduced Duration. Points represent group means, and error bars are Standard Errors.

Comparisons across learning instructions in immediate groups

Our final analysis compared the present immediate results with those of Experiment 2 in Wang & Gennari (2019). Recall that in this previous work, there was no Title-induced bias in reproduced durations, but participants were incidentally exposed to the titles—they were asked to read them while studying the animations to encourage language-induced event conceptualisations. Participants in this prior study were tested at varying times of day (morning and early afternoon). However, the absence of any difference between the Morning and Evening groups suggests that the time of day has little effect on reproduced duration. Except for the learning instructions, all other aspects of the design and tasks were identical to the present study. Therefore, comparing the present results with those in Experiment 2 of Wang & Gennari (2019) should reveal both Title-induced reproduction biases and other potential differences in title reporting and reproduced duration associated with specifically studying the titles and animations.

To compare across learning instructions, the model included Instruction as a factor, where the *Study Title* instruction corresponds to the present Immediate groups and *Incidental Title* corresponds to Experiment 2 in Wang & Gennari (2019). We used the coding strategy described in Section 2 and exemplified in Table 3 for the verbal recall data. Note that this coding differs from that previously reported for Experiment 2 in Wang & Gennari (2019) because the present hypotheses concern full title reports rather than title familiarity.

The reproduced duration results comparing across Instructions (Study Title vs Incidental Title) are presented in Table 9 and plotted in Fig. 6. Table 9.1 and Fig. 6 indicate that the instructions to study the titles led to generally longer reproduced durations in the Study Title group than the Incidental Title group, hence, the main effect of Instructions, although the differences were smaller for the longer stimulus durations (7 to 9 sec). This interactive pattern indicates that studying the titles lengthened reproductions for most stimulus durations and shifted the overall mean reproduced duration upwards (see the model with all interactions in SR-Table 4 of the OSF Supplementary Results).

As expected from the task instructions, there were also differences in title reporting: The Study group was more likely to report the titles in their verbal reports, as shown in Table 9.2, and this likelihood was unaffected by total word counts. Taken together, the results suggest that studying the titles led to more title reporting and title-induced biases

Table 9
Results from comparisons across learning instructions.

1. Reproduced Duration			
Predictors	Estimates	CI	p
Stimulus duration	0.76	0.59 – 0.92	**<0.001**
Stimulus segments	0.21	0.11 – 0.32	**<0.001**
Title (Slow vs Fast)	0.04	−0.01 – 0.10	0.101
Instructions (Study vs Incidental Title)	0.28	0.13 – 0.44	**<0.001**
Title * Instructions	0.05	−0.00 – 0.09	0.055
Random Effects			
σ^2	1.65		
$\tau_{00\ \mathrm{SubjID}}$	1.25		
$\tau_{00\ \mathrm{itemN}}$	0.11		
$\tau_{11\ \mathrm{SubjID.scale(duration)}}$	0.16		
$\tau_{11\ \mathrm{itemN.Title}}$	0		
$\rho_{01\ \mathrm{SubjID}}$	0.69		
$\rho_{01\ \mathrm{itemN}}$	−0.17		
ICC	0.48		
N $_{\mathrm{SubjID}}$	176		
N $_{\mathrm{itemN}}$	21		
Observations	3530		
Marginal R^2 / Conditional R^2	0.214 / 0.592		
2. Title reporting			
Predictors	***Odds Ratios***	***CI***	***p***
(Intercept)	0.52	0.30 – 0.91	**0.022**
Total word count	0.88	0.75 – 1.03	0.121
Instructions (Study vs Incidental)	1.71	1.07 – 2.73	**0.025**
Random Effects			
σ^2	3.29		
$\tau_{00\ \mathrm{SubjID}}$	1.37		
$\tau_{00\ \mathrm{itemN}}$	0.49		
ICC	7.14		
N $_{\mathrm{SubjID}}$	0.41		
N $_{\mathrm{itemN}}$	0.75		
Observations	173		
Marginal R^2 / Conditional R^2	21		
Groups	**Predicted Probabilties**	CI	
Study Title	0.47	.33 –.62	
Incidental Title	0.21	.12 –.34	

and introduced an additional temporal bias towards lengthening. This finding provides clues to qualify the lengthening effect observed earlier in the Sleep group, as discussed below.

Discussion

The present study investigated how learned videos titles conveying event schemas modulate temporal event memory before and after sleep. We specifically asked whether studying the titles with the animations would elicit title-induced temporal biases and whether sleep-dependent consolidation would additionally modulate temporal recollection. We found that studying video titles led to title biases in reproduced durations in all groups except the Wake group: On average, animations with titles implying fast speed were reproduced as shorter than those with titles implying slow speed. This finding contrasted with previous results showing that incidental title learning did not lead to title-induced biases (Wang & Gennari, 2019). All groups also showed some evidence of title reporting at recall, but the Sleep group was more likely than the Wake group to initiate verbal recall with the title, suggesting more robust title-video integration after sleep. Comparisons across learning instructions also indicated that incidental title learning led to lower rates of full title reporting at recall. These findings are consistent with the hypothesis that specifically learning the video titles would modulate reproduced duration, except when delay conditions interfere with recollection or consolidation, as in the Wake group.

We also found that episodic reproductions in all groups were correlated with the number of unique meaningful words reported, suggesting that despite the intervention of delays, duration reproductions remained linked to the episodic representation in similar ways as immediate reproductions. Nevertheless, the Sleep group systematically lengthened duration reproductions more than any other group (Fig. 4), suggesting a distinctive shift in the memory representation underlying duration reproduction. The convergence of intentional title learning, sleep-dependent consolidation, and title meanings may have induced stronger semantic relationships between title and video memories, leading to combined representations more strongly assimilated into the title schemas. This possibility is consistent with the comparison between incidental vs intentional title learning: intentional title learning lengthened the mean duration reproduction to some extent, suggesting that repeated title study modulates immediate event representations beyond implied speed differences (Fig. 6). Thus, sleep-dependent consolidation deepened the semantic relationship between video and title memories, leading to more prominent lengthening.

An unexpected result in this study was that the number of unique meaningful words reported in the Immediate groups was significantly smaller than in the Delay groups. Immediate memories are generally considered the most vivid and accurate compared to varying delay

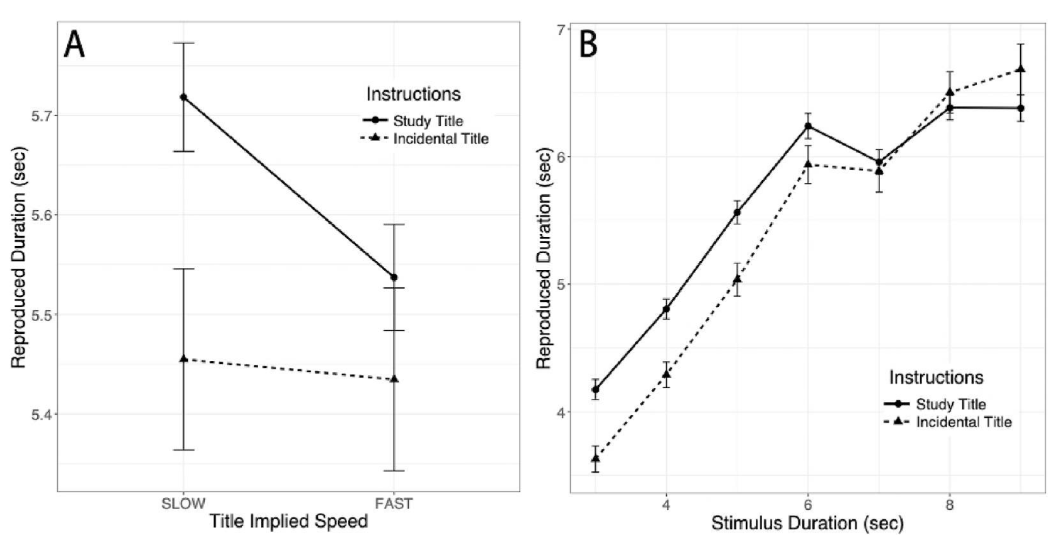

Fig. 6. Panel A shows Reproduced Durations as a function of Title and Learning Instructions. Panel B shows Reproduced Duration as a function of stimulus duration and Instructions. Dotted lines represent the group incidentally exposed to the Titles. Error bars represent Standard Error, and lines represent means.

conditions (Sekeres et al., 2018; Wixted & Ebbesen, 1991). Immediate reproductions were also the most accurate, i.e., the closest to stimulus duration, despite expected inaccuracy (Fig. 4A and 5A). Thus, shorter verbal reports in Immediate groups are unlikely to reflect poorer recollection. Instead, these groups were less likely to report all recalled details due to differences in session structure. The Immediate group completed learning and testing within the same session, lasting about an hour. In contrast, Delay groups performed these tasks separately and with plenty of allocated time. This setup may have prompted participants to limit their verbal recall task because, unlike mental replays, this task was more tiresome and the longest. Participants may have omitted typing everything they recalled and instead succinctly reported what happened in the video, often using the titles.

Interestingly, the influence of session structure on meaningful word counts did not prevent immediate groups from reporting the titles, as these did not significantly differ from those after sleep (see Table 8, Immediate vs Delay contrast). However, a sceptic may argue that the observed numerical difference is due to the session structure. This possibility is at odds with our title report measure. Recall that we have motivated this measure on the pairing strength between the visual cue and the title (see Section 2.5, *Title reporting*), but if participants strategically omitted title reports, e.g., to save time and effort, strategic factors may confound interpretation. Although we cannot entirely exclude this possibility, the Wake and Sleep groups were tested in equivalent session structures. Moreover, the ample difference between these groups (.36 vs.67) is unlikely to solely depend on task-induced strategies.

Overall, our results suggest that well-learned event labels can introduce immediate temporal biases and incipient lengthening that may be further modulated during sleep. Title-induced biases and lengthening co-occurred with higher title reporting after sleep but not wakefulness, but temporal reproductions maintained a general relation to stimulus duration and verbal recall, suggesting that they were grounded in memory traces. These modulations have implications for understanding the relationship between language and event memory.

Verbal labels and temporal event memory after learning

Previous research on the role of labels in event or object memory has often compared verbal and non-verbal tasks, e.g., object naming vs perception (Carmichael et al., 1932; Lupyan, 2008; Richler et al., 2013). In such studies, task requirements rather than language per se may direct attention to different stimulus aspects. In contrast, the present study kept task demands constant (studying titles and videos) and varied the accompanying labels, as in Feist and Gentner (2007). This design allows examining meaning rather than task modulations on subsequent memory performance. The experimental design and other differences with previous research, such as the stimulus characteristics, learning task, and the type of memory examined (recognition vs temporal memory), make comparisons and generalisations across studies difficult. For example, reproduction tasks are likely governed by different principles from those operating in recognition memory (Hanawalt, 1937). Nevertheless, the present results offer valuable insights into the interplay between verbal labels and dynamic visual events.

Our results suggest that despite typically inaccurate temporal recollection, the influence of labels is found when the association between titles and videos is robust and deeper. Intentionally learning the titles and animations elicited title biases and higher rates of title reporting than incidental title learning. We hypothesised that studying the titles and videos would deepen their semantic relation because their shared features would be foregrounded, leading to more semantic links or more elaborated semantic relations (Craik & Tulving, 1975; Einstein & Hunt, 1980; Staresina, Gray, & Davachi, 2009; Anderson & Reder, 1979). The repeated stimulus presentation in our study provided ample opportunity to encode the title's semantic features and rehearse its paring with the video, promoting the representation of the geometric figures as possible real-world objects and motion events. This integrated representation nudged duration reproductions during episodic reproductions towards the title meanings. This view is broadly consistent with interactive encoding accounts arguing for meaning modulations on the encoding of visual stimuli (Feist & Gentner, 2007; Lupyan, 2008).

However, our present and previous results are inconsistent with an encoding account arguing for spontaneous biases like those resulting from object labelling or event descriptions because repeated incidental title exposure did not elicit detectable title biases. Some researchers have argued that verbal stimulus encoding may lead to memory consequences because the verbalisation process requires linguistic analysis of the stimuli (Slobin, 1996). Although our study did not involve self-generated descriptions, stimulus titles are generally expected to relate to stimulus content (Bransford & Johnson, 1972; Huff & Schwan, 2008). Thus, contrasting labels do not necessarily elicit memory biases, even when they help event understanding.

One possible factor explaining the contrast between incidental and intentional title learning is the depth of title processing (Craik & Tulving, 1975). In incidental title understanding, participants may have understood the titles superficially and did not fully process the implied motion and speed features. Participants may have learned that moving figures represented real-world objects but did not link their motion features to the event dynamics. This possibility is consistent with additional exploratory analyses of verbal recall: Unlike title recall, the likelihood of using title nouns to refer to the geometric figures (e.g., *cadets* vs *athletes*) did not differ significantly across incidental and intentional title learning, suggesting that in both conditions, participants were familiar with the titles. However, only with intentional learning, the title event, e.g., *cadets crawling* vs *athletes racing,* was fully integrated with the dynamic stimuli.

Another factor promoting title biases on temporal memory is strengthening the video-title pairings and their semantic links. As in paired-associates learning, the joint encoding and rehearsal of videos and titles across several presentations promote semantic elaborations and links between the two items. The depth of this encoding facilitates the retrieval of one item when cued by the other (Craik & Tulving, 1975; Einstein & Hunt, 1980; Silverstein & Marshall, 1968; Tulving & Pearlstone, 1966; Anderson & Reder, 1979). This facilitation motivated the analysis of title reporting at the trial start: strongly linked items in memory cue each other at retrieval. The congruent semantic relation between titles and videos likely helped semantic elaboration and may have contributed to the temporal biases observed. However, it remains to be seen whether the intentional learning of semantically unrelated or incongruent title- video pairs would modulate temporal reproductions.

In sum, the results of the immediate groups demonstrate that joint intentional learning of visual and verbal information can bias temporal memory, leading to representations combining visual and verbal features. This combination was manifested in the concurrent influence of verbal labels, stimulus duration, and the number of event details recalled on episodic reproductions. These results thus suggest that encoding tasks encouraging deeper semantic processing can result in integrated conceptual and visual event representations.

Labels and temporal memory after delays

The results indicate that the role of verbal labels varied depending on delay processes. The Wake group showed poorer title recollection than the Sleep group and no detectable title biases in episodic reproductions. These results likely stem from the interference of wakeful activities with memory consolidation and/or retrieval processes, impairing recollection (Wixted, 2004, 2005; Berres & Erdfelder, 2021; Dewar et al., 2012). The forgetting of some title features and, possibly, some event details led to degraded and weaker stimulus memories that did not reliably include title-related motion features. This forgetting likely increased uncertainty at the point of reproductions, resulting in skewed and less accurate reproductions relative to immediate groups (see Fig. 4). This possibility is consistent with previous findings showing that poorer stimulus recall is associated with similar temporal skews. For example, comparing episodic reproductions after studying the stimuli once vs three times led to a similarly skewed pattern as that of the Wake group (Wang & Gennari, 2019).

In contrast, memory representations after sleep were strengthened compared to wakefulness, leading to more robust title-video integration. Many studies indeed argue that sleep-dependent consolidation promotes the integration of episodic details with prior schemas (Berres & Erdfelder, 2021; Feld & Born, 2017; Hennies et al., 2016; Klinzing et al., 2019; Lau et al., 2010, 2011; Stickgold & Walker, 2013). In the present study, the labels supplied the event schemas with which dynamic episodic details were combined. Consequently, sleep-dependent consolidation enabled better detail and title recall than wakefulness and promoted deeper title-video semantic integration. As in immediate groups, this semantic integration biased reproductions towards the title meanings, resulting in differences between slow and fast titles. Unlike any other group, however, episodic reproductions after sleep were systematically longer, suggesting an additional and distinctive contribution of sleep to the encoded relationship between videos and titles.

This additional contribution likely reflects deeper title-video integration than those soon after learning. After learning, visual stimulus features are more vivid and accessible than after longer delays, even if the titles were well-learned. During sleep, memory consolidation mechanisms may transform recently acquired associations through neuronal replay, strengthening connections between them and conceptual/semantic memory networks (Dudai, 2012; Klinzing et al., 2019). Deeper integration in intentional vs incidental title learning also shifted the mean reproduced duration, suggesting conceptual modulations towards lengthening even before sleep. Overnight consolidation then strengthened and further merged the relationship between the titles' semantic features and the video memory, leading to stronger conceptual influences on episodic reproductions.

A deeper title-video integration explains why conceptual influences lead to systematically lengthened reproductions rather than some other pattern. Attributing real-world characteristics to the video motion segments leads to longer segment reproductions than viewing them as moving figures. For example, mentally replaying an ambulance travelling to the hospital in Fig. 1 involves turning city corners, picking someone up, dropping the passenger off, etc. In contrast, mentally reenacting recently seen figures moving on a screen is less constrained by familiar action knowledge. Thus, post-sleep reproductions were more strongly assimilated into the label's meaning. Nevertheless, they remained grounded in the event details recalled, as indicated by correlations with stimulus segments, stimulus durations and verbal recall (Figs. 4 and 5). Thus, post-sleep memory representations contained both conceptual and episodic features that were more deeply integrated, leading to further distortions in episodic reproductions.

The present results demonstrate that encoding verbal labels and videos modulates temporal event recollection after sleep to a greater extent than after learning or a period of wakefulness. This finding suggests that verbal labels and sleep-dependent consolidation may introduce additional memory biases absent immediately after learning, thus highlighting the potential role of language in shaping cognitive representations more generally.

Conclusions

The present study aimed to determine whether and how learned verbal event labels modulate temporal event memory before and after delays, thus illuminating the mechanisms underpinning the interaction of language and memory. We found that joint learning of verbal labels and videos elicited conceptual biases in temporal recollection soon after learning, leading to longer or shorter mental event reproductions consistent with the labels' implied speed. These reproductions were also longer to some extent compared to previous studies involving incidental exposure to verbal labels. Temporal recollection after a period of sleep but not wakefulness also displayed speed-related biases, but post-sleep mental reproductions were markedly longer than at any other recollection point. The joint learning of labels and videos and sleep-dependent consolidation thus promoted different degrees of semantic integration before and after sleep. These results suggest that the degree of semantic integration between visual and verbal stimuli is a critical factor in shaping temporal biases and, more generally, in determining the role of verbal labels in event learning and memory.

CRediT authorship contribution statement

Yaqi Wang: Writing – original draft, Visualization, Validation, Project administration, Investigation, Formal analysis, Data curation. **M. Gareth Gaskell:** Writing – review & editing, Conceptualization. **Silvia P. Gennari:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

https://doi.org/10.17605/OSF.IO/K8YXA.

References

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*(2), 203–231.

Aly, M., & Moscovitch, M. (2010). The effects of sleep on episodic memory in older and younger adults. *Memory, 18*(3), 327–334. https://doi.org/10.1080/09658211003601548

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak, & F. I. M. Craik (Eds.), *Levels of Processing in Human Memory* (pp. 385–404). Erlbaum.

Arnold, A. E. G. F., Iaria, G., & Ekstrom, A. D. (2016). Mental simulation of routes during navigation involves adaptive temporal compression. *Cognition, 157*, 14–23. https://doi.org/10.1016/j.cognition.2016.08.009

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berres, S., & Erdfelder, E. (2021). The sleep benefit in Episodic memory: An integrative review and a meta-analysis. *Psychological Bulletin*. https://doi.org/10.1037/bul0000350.supp

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*(1), 1–47.

Bock, J. K. (1987a). An effect of the accessibility of word forms on sentence structure. *Journal of Memory and Language, 26*, 119–137.

Bock, J. K. (1987b). Coordinating words and syntax in speech plans. In A. W. Ellis (Ed.), *Progress in the Psychology of Language* (pp. 337–390). Erlbaum.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*, 47–67.

Bonasia, K., Blommesteyn, J., & Moscovitch, M. (2016). Memory and navigation: Compression of space varies with route length and turns. *Hippocampus, 26*, 9–12. https://doi.org/10.1002/hipo.22539

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 717–726. https://doi.org/10.1016/S0022-5371(72)80006-9

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*, 193–213.

Carmichael, L., Hogan, H. P., & Walter, A. A. (1932). An experimental study of the effect of language on the reproduction of visually perceived forms. *Journal of Experimental Psychology, 15*, 73–86.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

D'Argembeau, A., Jeunehomme, O., & Stawarczyk, D. (2021). Slices of the past: How events are temporally compressed in episodic memory. *Memory*, 1–6. https://doi.org/10.1080/09658211.2021.1896737

Dewar, M., Alber, J., Butler, C., Cowan, N., & Della Sala, S. (2012). Brief wakeful resting boosts new memories over the long term. *Psychological Science, 23*(9), 955–960. https://doi.org/10.1177/0956797612441220
Dudai, Y. (2012). The restless engram: Consolidations never end. *Annual Review of Neuroscience, 35*, 227–247. https://doi.org/10.1146/annurev-neuro-062111-150500
Dumay, N., & Gaskell, M. C. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35–39.
Einstein, G. O., & Hunt, R. R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 588.
Ekstrand, B. (1977). The effect of sleep on human long-term memory. In R. Drucker-Colin (Ed.), *Neurobiology of Sleep and Memory* (pp. 419–438). Academic.
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
Feist, M. I., & Gentner, D. (2007). Spatial language influences memory for spatial scenes. *Memory and Cognition, 35*(2), 283–296.
Feld, G. B., & Born, J. (2017). Sculpting memory during sleep: Concurrent consolidation and forgetting. *Current Opinion in Neurobiology, 44*, 20–27. https://doi.org/10.1016/j.conb.2017.02.012
Folkard, S. (1979). Time of day and level of processing. *Memory & Cognition, 7*(4), 247–252. https://doi.org/10.3758/BF03197596
Gilboa, A., & Marlatte, H. (2017). Neurobiology of schemas and schema-mediated memory. *Trends in Cognitive Sciences, 21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013
Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*(3), 561–582. https://doi.org/10.3758/APP
Hanawalt, N. G. (1937). Memory trace for figures in recall and recognition. *Archives of Psychology, 216*, 1–81.
Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. *Journal of Experimental Psychology. General, 118*(2), 136–147. http://www.ncbi.nlm.nih.gov/pubmed/2525593.
Hemmer, P., & Steyvers, M. (2009a). A bayesian account of reconstructive memory. *Topics in Cognitive Science, 1*(1), 189–202. https://doi.org/10.1111/j.1756-8765.2008.01010.x
Hemmer, P., & Steyvers, M. (2009b). Integrating episodic memories and prior knowledge at multiple levels of abstraction. *Psychonomic Bulletin and Review, 16*(1), 80–87. https://doi.org/10.3758/PBR.16.1.80
Hennies, N., Lambon Ralph, M. A., Kempkes, M., Cousins, J. N., & Lewis, P. A. (2016). Sleep spindle density predicts the effect of prior knowledge on memory Consolidation. *The Journal of Neuroscience, 36*(13), 3799–3810. https://doi.org/10.1523/JNEUROSCI.3162-15.2016
Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: A new approach. *Psychophysiology, 10*, 431–436.
Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*, 97–110.
Huff, M., & Schwan, S. (2008). Verbalizing events: Overshadowing or facilitation? *Memory and Cognition, 36*(2), 392–402. https://doi.org/10.3758/MC.36.2.392
Huttenlocher, J., Hedges, L. V., & Bradburn, N. M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(2), 196–213. https://doi.org/10.1037/0278-7393.16.2.196
Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and Particulars: Prototype Effects in Estimating Spatial Location. *Psychological Review, 98*(3), 352–376.
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
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., Folville, A., Stawarczyk, D., Van der Linden, M., & D'Argembeau, A. (2017). Temporal compression in episodic memory for real-life events. *Memory, 26*(6), 1–12. https://doi.org/10.1080/09658211.2017.1406120
Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience, 22*(October). https://doi.org/10.1038/s41593-019-0467-3
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
Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. *Sleep Medicine Reviews, 18*(6), 531–541. https://doi.org/10.1016/j.smrv.2014.03.005
Lau, H., Alger, S. E., & Fishbein, W. (2011). Relational memory: A daytime nap facilitates the abstraction of general concepts. *PLoS ONE, 6*(11). https://doi.org/10.1371/journal.pone.0027139
Lau, H., Tucker, M. A., & Fishbein, W. (2010). Daytime napping: Effects on human direct associative and relational memory. *Neurobiology of Learning and Memory, 93*(4), 554–560. https://doi.org/10.1016/j.nlm.2010.02.003
Lichtenstein, E. H., & Brewer, W. F. (1980). Memory for goal-directed events. *Cognitive Psychology, 12*(3), 412–445. https://doi.org/10.1016/0010-0285(80)90015-8
Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361–366. https://doi.org/10.1101/lm.94705
Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobil destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior, 13*, 585–589.
Lupyan, G. (2008). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General, 137*(2), 348–369. https://doi.org/10.1037/0096-3445.137.2.348
Matthews, W. J., & Meck, W. H. (2016). Temporal cognition: Connecting subjective time to perception, attention, and memory. *Psychological Bulletin, 142*(8), 865–907. https://doi.org/10.1037/bul0000045
McDonald, J. L., Bock, J. K., & Kelly, M. H. (1993). Word and world order: Semantics, phonological, and metrical determinants of serial position. *Cognitive Psychology, 25*, 188–230.
Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology, 12*(5), 436–450. https://doi.org/10.1016/0022-1031(76)90076-7
Oberauer, K., & Greve, W. (2022). Intentional remembering and intentional forgetting in working and long-term memory. *Journal of Experimental Psychology: General, 151*(3), 513–541. https://doi.org/10.1037/xge0001106
Popov, V., & Dames, H. (2022). Intent matters: Resolving the intentional versus incidental Learning Paradox in Episodic long-term memory. *Journal of Experimental Psychology: General, 152*(1), 268–300. https://doi.org/10.1037/xge0001272
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
Rasch, B. H., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews, 96*(2), 681–766. https://doi.org/10.1152/Physrev.00032.2012
Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences, 13*(10), 439–446. https://doi.org/10.1016/j.tics.2009.07.001
Richler, J. J., Gauthier, I., & Palmeri, T. J. (2011). Automaticity of basic-level categorization accounts for labeling effects in visual recognition memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 37*(6), 1579–1587. https://doi.org/10.1037/a0024347
Richler, J. J., Palmeri, T. J., & Gauthier, I. (2013). How does using object names influence visual recognition memory? *Journal of Memory and Language, 68*(1), 10–25. https://doi.org/10.1016/j.jml.2012.09.001
Roy, M. M., & Christenfeld, N. J. S. (2008). Effect of task length on remembered and predicted duration. *Psychonomic Bulletin and Review, 15*(1), 202–207. https://doi.org/10.3758/PBR.15.1.202
Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1481), 773–786. https://doi.org/10.1098/rstb.2007.2087
Schacter, D. L., & Addis, D. R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philosophical Transactions of the Royal Society B, 364*(1521), 1245–1253. https://doi.org/10.1098/rstb.2008.0308
Sekeres, M. J., Winocur, G., & Moscovitch, M. (2018). The hippocampus and related neocortical structures in memory transformation. *Neuroscience Letters, 680*(May), 39–53. https://doi.org/10.1016/j.neulet.2018.05.006
Shank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Erlbaum.
Silverstein, A., & Marshall, A. (1968). Incidental vs. intentional paired-associate Learning. *The American Journal of Psychology, 81*(3), 415. https://doi.org/10.2307/1420639
Slobin, D. I. (1996). From "thought and language" to "thinking for speaking". In J. Gumperz, & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge University Press.
Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebral Cortex, 19*(5), 1198–1207. https://doi.org/10.1093/cercor/bhn165
Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience, 16*(2), 139–145. https://doi.org/10.1038/nn.3303
Tilley, A., & Warren, P. (1983). Retrieval from semantic memory at different times of day. In *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9* pp. 718–724). https://doi.org/10.1037/0278-7393.9.4.718
Tompary, A., & Thompson-Schill, S. L. (2021). Semantic influences on Episodic memory distortions. *Journal of Experimental Psychology: General, 150*(9), 1800–1824. https://doi.org/10.1037/xge0001017
Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2). Addison-Wesley.
Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 381–391. https://doi.org/10.1016/S0022-5371(66)80048-8
van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Cognitive Sciences, 35*(4), 211–219.
von Sobbe, L., Maienborn, C., Reiber, F., Scheifele, E., & Ulrich, R. (2021). Speed or duration? effects of implicit stimulus attributes on perceived duration. *Journal of Cognitive Psychology, 33*(8), 877–898. https://doi.org/10.1080/20445911.2021.1950736
Wang, Y., & Gennari, S. P. (2019). How language and event recall can shape memory for time. *Cognitive Psychology, 108*, 1–21. https://doi.org/S0010028518301610.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55*, 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Wixted, J. T. (2005). A Theory About Why We Forget What We Once Knew. *Current Directions in Psychological Science, 14*(1), 6–9.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science, 2*(6), 409–415. https://doi.org/10.1111/j.1467-9280.1991.tb00175.x

Wolford, G. (1971). Function of distinct associations for paired-associate performance. *Psychological Review, 78*, 303–313. https://doi.org/10.1037/h0031032

Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: Systems consolidation reconsidered. *Nature Reviews Neuroscience, 20*(6), 364–375. https://doi.org/10.1038/s41583-019-0150-4

Yuan, L., Novack, M., Uttal, D., & Franconeri, S. (2024). Language systematizes attention: How relational language enhances relational representation by guiding attention. *Cognition, 243*, Article 105671. https://doi.org/10.1016/J.COGNITION.2023.105671

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., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology. General, 138*(2), 307–327. https://doi.org/10.1037/a0015305

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