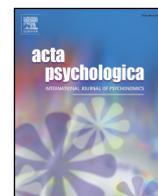




Contents lists available at ScienceDirect

Acta Psychologica

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

Representing time in language and memory: The role of similarity structure

Myrthe Faber, Silvia P. Gennari *

University of York, UK

ARTICLE INFO

Available online xxxx

PsycINFO codes:

2340 Cognitive processes
2343 Learning and memory
2700 Communication systems
2720 Linguistics

Keywords:

Event memory
Event duration
Similarity
Retrospective time
Reading
Duration estimation

ABSTRACT

Every day we read about or watch events in the world and can easily understand or remember how long they last. What aspects of an event are retained in memory? And how do we extract temporal information from our memory representations? These issues are central to human cognition, as they underlie a fundamental aspect of our mental life, namely our representation of time. This paper reviews previous language studies and reports a visual learning study indicating that properties of the events encoded in memory shape the representation of their duration. The evidence indicates that for a given event, the extent to which its associated properties or sub-components differ from one another modulates our representation of its duration. These properties include the similarity between sub-events and the similarity between the situational contexts in which an event occurs. We suggest that the diversity of representations that we associate with events in memory plays an important role in remembering and estimating the duration of experienced or described events.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

We represent events as taking time. If we watch someone cook dinner, we are able to reconstruct the time this event took when later remembering it. Similarly, when we read descriptions of events, we have a sense of how long those events may last. For example, upon reading *the police arrested the suspect* or alternatively, *the police held the suspect*, somehow we understand that the latter event was longer than the former. How do we represent the duration of events we have experienced or read about? What aspects of the events represented in memory contribute to our sense of how long they take?

Here we review some relevant findings in language research and report a memory study addressing these questions. We build on previous theories within the psychology of time literature, in particular, those concerning memory-based time estimations (Block, 1982; Block & Zakay, 1997; Boltz, 2005; Grondin, 2010). When participants do not know beforehand that they will be asked for duration estimations, they will have to retrospectively reconstruct duration based on their memory of what happened (Block, 1974; Block & Reed, 1978; Ornstein, 1969). Previous studies have shown effects of memory content and processing context variability on these duration estimations. For example, Ornstein (1969) asked participants (one trial per

participant) to judge the relative length of two intervals that were in fact of the same clock duration. One instance of this interval was filled with few tones, whereas the other was filled with several tones. It was found that participants judged the interval filled with more tones as being longer than the other. Likewise, Block (1982) manipulated contextual changes by altering the physical environment or the task instructions under which the stimuli to be compared were first processed (e.g., attending to structural or semantic characteristics of the stimuli, or changing the room in which they were processed). Subsequent duration judgments revealed that contextual changes led to overestimation of duration. Based on findings of this sort, it has been proposed that stimulus complexity, the variability of contextual changes associated with the stimuli, and generally, the amount of information gleaned from the stimulus, determines retrospective duration estimations.

Retrospective estimations are similar to those elicited by language in that they are based on memory representations. Representations elicited by language rely on semantic memory representations built over time from our experience of events and the mapping of words onto event representations. Retrospective representations of events on the other hand, rely on the memory representations built from previous experience of the events. Such representations may emerge from a single experience of an event, as in most retrospective studies, but may also emerge from repeated experiences that are learned and encoded over time into memory schemas. We explore this link here, and attempt to establish some relationships between long-term representations in semantic memory and those that emerge from learning,

* Corresponding author at: Department of Psychology, University of York, Heslington, York YO10 5DD, UK.

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

on the assumption that learning mechanisms that initially apply when we first learn about events during the course of our lives are similar to those operating in learning stimuli for which we do not have already established memory representations. Much research in statistical learning has indeed reported parallel mechanisms across adults and children in extracting schemas or generalizations from regularities in the stimulus (Gómez & Gerken, 2000; Saffran, Johnson, Aslin, & Newport, 1999).

We therefore argue that some aspects of event representations in semantic memory, which help us to estimate their duration when they are described by language, also play a role when estimating duration of recently learned visual events. In particular, we argue that the similarity structure of the memory representations modulates our duration estimates in both of these domains, language and visual memory. Of course, pointing out some parallels across these domains does not entail that there will not be important differences. Indeed, it is almost certain that schemas in semantic memory have many more properties than those emerging from a few exposures to novel visual events, given the multidimensional nature of events, the multiple contexts in which they can occur and the hierarchical structure that they afford (Zacks, Speer, Swallow, Braver, & Reynolds, 2007), all aspects that likely need longer periods of exposure and memory consolidation than those present in a single learning study. Yet, such parallels may turn up to be illuminating and guide further research.

1.1. The puzzle

Research in mental models has suggested that as we read discourse or sentences, we build a mental model of the temporal relationships between events that resembles properties of the situations described (Kelter, Kaup, & Claus, 2004; Zwaan & Radvansky, 1998; Zwaan, Radvansky, Hilliard, & Curiel, 1998). For example, longer temporal distances between events have been shown to take longer to process (Gennari, 2004; Zwaan, 1996) and information from temporally more distant events in the narrative takes longer to retrieve (Kelter et al., 2004). This research suggests that mental representations of temporal distance are analogical with the characteristics of the world described. However, to what extent can our mental representations be analogical with actual time? The representations that we entertain during reading or remembering are clearly not a copy or a re-enactment of the actual events described or experienced, as for instance longer described duration does not lead to a linear increase in processing cost: it does not take a lifetime to process a sentence like *John searched for the right woman all his life*. Rather, these memory representations are somehow temporally compressed, activating salient or typical characteristics of the events being talked about. The question therefore arises what aspects of the events are retained in memory and activated during processing, either when retrieving word meanings and composing the representation of the event described or when retrieving an event from memory.

1.2. Some observations from language research

In an attempt to answer the question above, Marta Coll-Florit and one of the present authors (Coll-Florit & Gennari, 2011) investigated the processing of event descriptions that differed in the duration attributed to the referred events, and correlated their processing times with measures that characterized semantic properties of the events in long-term memory. In a semantic decision task (where participants were asked to indicate whether stimulus phrases made sense), we directly compared the response times to phrases such as *to owe money* and *to lose money*, which differed in the duration of the event described. We found that indeed verb phrases that refer to durative events (*to owe money*) took longer to process than phrases referring to non-durative events (*to lose money*). This suggested that longer events activate event knowledge that is more costly to access or retrieve. Similar results were found in a second study with Spanish verbs that undergo changes in interpretation (or sense) from longer to shorter events as a function

of imperfective or perfective morphology (e.g., *conocía/conoció* 'knew/learned'). Interestingly, processing times of these verbs and verb phrases positively correlated with the duration attributed to the referred events in an arbitrary 1–7 scale.

What drives the differences in processing time in these studies? And why does event duration correlate with processing times? To begin to address these questions, we conducted a free association study often used in memory research to establish the words that come to mind upon reading our stimulus verbs and phrases (Coll-Florit & Gennari, 2011). Analyses of these associations showed that durative events elicited associations of a more diverse nature than non-durative events: a non-durative event brings to mind mostly other events and objects (e.g., for *losing money*, associates would be *money, gambling, mistake, drop*), whereas durative events elicit a variety of semantic types including states and properties as well as other events and objects (for *owing money*, associates would be *slow to pay, worried, loan, gratitude*). We also conducted a corpus study investigating the sentential contexts in which the verbs used in our phrasal study tend to occur. The verbal contexts in which verbs occur provide information about the real world situations in which the events referred to by the verbs occur, as they indicate the participants, objects, locations, etc., with which the events co-occur. The verbal contexts are thus a proxy for the properties of real world situation contexts in which events occur. We computed the Euclidean distance between these contexts, a measure indicating how similar the situation contexts are, thus providing a measure of how diverse the verbs' memory associations built over time through experience of the events are. We found that the similarity of the non-durative contexts is larger than that of the durative ones, consistent with the idea that the contexts in which durative events occur are more variable and thus, less similar to those of all other verbs. Importantly, the Euclidean distance of a verb in semantic space and a measure of variability between memory associations from the association study correlated with both duration estimates for the stimuli and their processing times. These results demonstrate a link between the diversity of situational contexts and memory associates of event descriptions with both estimated duration and processing times, suggesting the possibility that these two behavioral measures are driven by properties of the representations established in semantic memory.

The results described here have been extended to different sentential stimuli for which the event described does not depend on lexical meaning alone. In fact, Gitte Joergensen and one of the present authors (Joergensen & Gennari, 2013) compared the interpretation of the same action verbs in different contexts that implied long and short durations (e.g., *to alter a dress for one hour vs. three hours*). We found that the time it takes to retrieve the event described in a memory task depended on duration and also correlated with the diversity or variability of the memory associations generated by the discourse situation. This study, together with the study manipulating Spanish morphology mentioned above, addresses a concern in many previous studies, namely, that the interpretation of duration is often correlated with different verb classes or lexical meaning. However, how long we interpret an event to be depends on its context, and this interpretation does not always correlate with verb classes, and instead depends on whatever cues are available in the context, thus, highlighting the orthogonal and distinctive nature of a given interpretation based on event knowledge and grammatical features (Metusalem et al., 2012). Taken together, these results suggest that the diversity of associations linked to events in semantic memory may be responsible for their processing cost. Although the definitive answer for why this might be the case awaits further research, our results suggest that the cost of retrieving event knowledge or schemas may result from the effort of collecting diverse properties and associations from memory, i.e., properties that come from different semantic domains and occur in diverse, dissimilar contexts.

Drawing a parallel to event memory, these results bring up the possibility that the diversity of associations for events in memory may be related to the duration attributed to them. Why should this be the

case? Previous memory-based duration judgments suggested that contextual changes and the amount of information encoded during stimulus processing modulate duration estimations. Because providing a duration estimate for an event entails retrieving the memory representation of that event, it might be the case that retrieving more and more variable information leads to longer estimations, even for events of the same clock duration. Thus, in Ornstein's experiments, more complex stimuli (e.g., larger number of tones, or more complex pictures) elicited longer duration estimates than simpler stimuli and in Block and colleagues' experiments, changes in processing contexts also elicited increased duration estimates relative to those without changes. Hence, events that are more complex in nature and are associated with variable contexts appear to take more space in semantic memory, and would therefore seem longer as more information needs to be retrieved from semantic memory.

Memory and learning research unrelated to the psychology of time literature supports this possibility. Similarity between events, and particularly, the repetition of similar events in variable contexts, leads to better, more efficient, memory encoding (Belezza & Young, 1989; Melton, 1967). Over subsequent exposures, similar events tend to be chunked into one schema (Brady, Konkle, & Alvarez, 2009), but dissimilar ones tend to be stored separately. Thus, chunking information and similarity are closely linked. Similarity has also been extensively studied in categorization, which indicates that clusters of features that are correlated in our experience play an important role in categorization and inductive learning (Sloutsky, 2003). More generally, the extraction of regularities and patterns of co-variation in our experience is essential for learning and the development of memory representations, which over time acquire a schematic or more abstract nature, as suggested by the statistical learning literature. Such regularities ultimately depend on the similarity between different instances of a same element/category in our experience and the similarity of the contexts in which they occur (Gómez & Gerken, 2000). The more similarities we perceive in our experience, the more information can be chunked together in patterns or schemas. It therefore appears that fewer similarities within the information encoded in memory engender more stored information, and this property of the memory representations might be linked, as previous retrospective time estimation studies suggest, to longer duration estimates.

2. Learning and encoding of visual events

In this study, we aimed to explore the claim that more stored information, and in particular, less similarity between the events perceived, will lead to longer duration estimates, despite the clock duration of the events remaining constant. To this end, we used a paired-associate learning paradigm in which novel stimulus animations were first encoded in memory. After learning and unbeknown to participants, duration estimations for each animation relative to that of all other animations were obtained. The animations depicted various geometric shapes moving, changing or causing changes. These novel events did not bring to mind typical human actions but were designed to contain internal structures attested in the real world, e.g., events with repetitive sub-events (hammering) or variable sub-events (building). Participants in the study learned the animations together with a still frame used as retrieval cue in subsequent recognition memory and duration estimation tasks.

We manipulated similarity and the complexity of the animations. We created a set of 15 animation triads varying in clock duration, but within a triad, the three animations lasted the same clock duration. Triad members were arranged in three conditions aimed at replicating complexity effects such as those studied by Ornstein, and more importantly, demonstrating an effect of similarity. Our focus here is on the similarity manipulation because it is more relevant for the language research discussed above, in particular, the role of semantic diversity in memory representations, although we report the full set of results for completeness. As shown in Fig. 1, the *basic condition* contained an

event sequence with a repeating or stable motion of a shape, the *numerous or complex condition* contained a sequence with a repeating change (e.g., displacement or color change) added onto the basic motion, and the *dissimilar condition* contained a sequence like the numerous one but with dissimilar changes (e.g., displacement and disappearance, changes into different colors). Note that the only difference between the numerous and dissimilar conditions is that different events occur in the sequence, but the number of things happening stays constant. These two conditions are therefore critical here to evaluate the influence of similarity. A pre-test questionnaire study established that indeed the numerous and the dissimilar conditions differed in the similarity between their sub-events, with the numerous condition having more similar sub-events than the dissimilar conditions.

We therefore predicted that when the similarity between sub-events decreases, duration estimates should increase, even when clock duration remains constant, because dissimilarity should lead to more stored information, and the more we recollect about an event, the longer it seems. Moreover, because the memory representations encoded for the stimulus animations should contain more or less information as a function of similarity, we expected that latencies in the memory task should show a similar pattern.

3. Methods

3.1. Participants

Fifty-two native English speakers, students from the University of York, participated in this experiment. Ten participants were excluded from the analysis because they had poor memory performance in one of the conditions (recognition accuracy $\leq 60\%$), and thus contributed few data points to the condition mean given the number of stimuli. Recognition accuracy was matched across conditions to avoid guesses (i.e., to make sure participants indeed had a mental representation of the animations studied) when providing duration estimations (basic: $M = 98\%$, $SD = 7\%$; numerous: $M = 94\%$, $SD = 10\%$; dissimilar: $M = 94\%$, $SD = 9\%$). Friedman's test indicated no significant differences across conditions.

3.2. Stimuli

Fifteen triads were created with Adobe Flash software and contained a basic, numerous, and dissimilar version, as indicated above. All members of the triads had the same clock duration but across triads had different durations randomly varying between 3 and 16 s. Triads contained animations such as those exemplified in Fig. 1. The basic version in a triad (see above) was systematically modified into the numerous version, which in turn was modified into the dissimilar one, keeping speed of motion constant across triad members. Across triads, shapes, motion and changes were visually different to prevent memory interference. For each triad, a single frame near the beginning of the animation, common to all triad members, was used as retrieval cue (Fig. 1, cue-frames). Foil frames were also extracted for the recognition memory task (see below).

3.3. Stimulus pre-test

A questionnaire study was conducted over the web to obtain similarity ratings. Animations were arranged in three lists (each containing one member of each triad but all three conditions across triads – Latin square design). Ten native English speakers were tested for each list. Means and standard deviations are displayed in Table 1.

The questionnaire asked participants to judge how similar the events within each animation were to one another on a scale of 1–7 (1 = not similar at all, 7 = very similar, using only integers). Examples were provided indicating the extreme points of the scale. The animations could be watched as many times as desired. A repeated measures ANOVA with items as random factor, condition as a repeated factor with three levels (basic, numerous and dissimilar) and

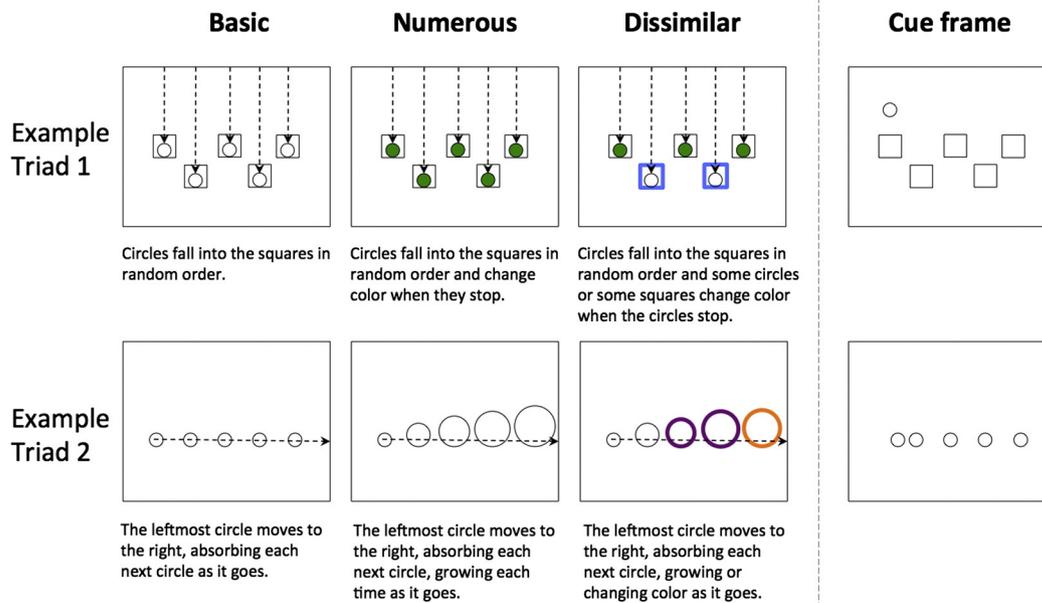


Fig. 1. Example of animation triads. Dotted lines indicate the motion path of the moving shape.

rating as a dependent variable indicated that there was a main effect of condition ($F(2, 28) = 19.33, p < .001, \eta_p^2 = .58$). Pair-wise comparisons showed that there was a significant difference between the numerous and the dissimilar condition ($t(14) = 3.80, p = .002$), but not between the basic and the numerous condition. (A by-subject analysis showed identical results $F(2, 18) = 50.72, p < .0001, t(9) < .0001$). This pattern of ratings confirms that the two conditions of interest here – the numerous and dissimilar conditions – indeed differ in perceived similarity, as intended. These results reflect the repetitive structure built into the basic and numerous conditions, but not the dissimilar one, which received very low ratings in comparison with the remaining conditions.

3.4. Design and procedure

Animations were arranged in three lists, each containing 15 animations (5 of each condition) but one animation per triad (Latin square design). Thus, participants only saw one member of each triad but several animations per condition. The experiment consisted of three tasks: a study task, a recognition task and a duration estimation task. In the study task, participants were instructed to study the content of the animations in association with the cue-frame for a subsequent memory task. They were told that the frames would be used to remind them of the whole animation in the later memory task. Each trial of the study task consisted of the presentation of a cue-frame for 2 s, followed by the corresponding animation. After each animation, participants pressed a key to move on to the next frame–animation pair. To facilitate learning of the animation content, the animations of a given list were presented three times, each time in a random order. In between cycles, participants were prompted to press a key to see and study the animations again.

After the study phase, participants performed a recognition memory task in which they were asked to indicate by pressing one of two keys on a computer keyboard whether a given cue-frame belonged to an animation that they had studied. As described above, 15 cue-frames (target-

frames) were obtained by extracting a frame from the beginning of an animation within each triad, so that this frame was common to all members of a triad (Fig. 1). If recognized correctly, these target-frames should elicit a YES-response. In order to balance the number of YES and NO responses, 15 foil-frames were also extracted from the animations. As the unfolding of events is not the same for all members of a triad, a middle or late frame from a triad member that was not initially studied was used as a foil. Thus, if the participant had studied the basic condition for a certain triad, the foil-frame would be obtained from the numerous or dissimilar condition and vice versa. This means that the foil-frames were visually similar to the studied animations, yet participants had to use their memory of the event sequences to correctly reject the foil and provide a NO-response. In each trial, a frame appeared on the screen until participants provided a response. Inter-trial times varied randomly between 500 and 3000 ms to avoid anticipation effects. The 15 target-frames and 15 foil-frames were presented in random order during the task. From this task, we analyzed response latencies to correctly accepted YES-responses. Recognition latencies longer than 2.5 standard deviations from the condition mean were considered outliers and therefore excluded from the data.

Finally, unbeknown to participants, they were asked to perform a final duration judgment task. In this task, participants judged the duration of the animations associated with the cue-frames on a scale of 1 to 7 using only the integers. “1” on the scale represented the shortest video seen in the study phase (*very short*), and 7 indicated the longest video (*very long*) as compared to the other animations that they have seen. When probed by the cue-frame, participants indicated their ratings by pressing a number key on the keyboard. To facilitate this judgment, an image of the 1–7 scale with the respective labels for the extreme values was displayed underneath the cue-frame. The trial terminated when participants provided a number between 1 and 7. Participants were instructed to use the whole scale. The cue-frames were presented in random order.

4. Results

4.1. Memory judgment latencies

We expected that if similarity structure determines encoding during learning and is then re-activated in cued-recognition, a difference between the numerous and the dissimilar condition should be observed in response latencies. Repeated measures ANOVAs with subjects or

Table 1
Mean similarity ratings per condition.

Event properties	Conditions		
	Basic	Numerous	Dissimilar
Similarity	5.71 (1.01)	5.23 (.82)	4.23 (.68)

Standard deviations are given in parentheses.

items as a random factor, experimental condition as a repeated factor (basic, numerous, dissimilar) and response latencies as a dependent variable indicated a main effect of condition ($F_1(2, 65) = 13.39, p < .001, \eta_p^2 = .246$ (Greenhouse–Geisser corrected); $F_2(2, 28) = 15.32, p < .001, \eta_p^2 = .523$). Importantly, the contrast between the numerous and dissimilar condition was significant ($t_1(41) = -3.19, p = .003; t_2(14) = -3.29, p = .005$), suggesting an effect of similarity. These results indicate that participants reactivated the content of the animations during recognition, with the dissimilar condition eliciting longer recognition times than the numerous condition. Means and standard errors are given in Fig. 2.

4.2. Duration ratings

To test the prediction that the dissimilar condition elicits longer estimates than the numerous condition despite clock duration remaining constant, we conducted repeated measures ANOVAs with subjects or items as a random factor, experimental condition as a repeated factor and duration ratings as a dependent variable. This analysis indicated a main effect of condition ($F_1(2, 82) = 9.30, p < .001, \eta_p^2 = .185; F_2(2, 28) = 5.74, p = .008, \eta_p^2 = .291$) and a significant difference between the numerous and dissimilar condition in the expected direction ($t_1(41) = -2.45, p = .02; t_2(14) = -2.40, p = .03$). Overall, these results suggest that the estimation of duration depends on the similarity structure people have encoded, with the dissimilar condition receiving higher estimates than the numerous condition.

4.3. Regression analysis

To evaluate the influence of similarity structure on duration estimations across all animations, we conducted by-item hierarchical multiple regressions. We examined the proportion of variance accounted for by similarity, over and above clock duration. We used a model with clock duration as a control predictor to account for the variation in actual duration across items (a factor that was controlled for in the categorical analyses above by the triad design). We found that adding similarity ratings to this regression model significantly increased the proportion of variance accounted (R^2) from .528 to .574 ($F_{\text{change}}(1, 42) = 4.56, p = .04$). Thus, sub-event similarity plays a role in duration estimation, over and above clock duration, suggesting that participants estimate duration from the event structure encoded in memory. Interestingly, memory judgment latencies were not correlated with clock duration, but they were correlated with similarity ratings ($r = -.43, p = .003$), with less similar sub-events leading to longer latencies. This is most

likely due to the reactivation of event characteristics in recognition tasks like ours, as suggested by memory studies (Yonelinas, 2001).

These results suggest that similarity structure is an important aspect of our memory representations, which modulates our reconstruction of the time passed: less similar sub-events lead to longer response latencies in recognition memory and longer duration estimates. Since less similarity between event components leads to more stored information, as suggested by our memory task and the contextual change hypothesis (Block, 1982), the more information has been stored in association with an event (animation), the longer the event is estimated to be.

5. General discussion

In the introduction, we asked what properties of events are retained in memory so that we can represent their temporal development in a compressed or schematic manner. The results reviewed and reported above suggest that the similarity structure encoded in memory (whether in episodic or semantic memory) may play a significant role in representing events and their duration, because this property determines the amount of information encoded and subsequently retrieved during processing and duration estimation. Thus, both descriptions of events with semantically diverse situational contexts and remembered event sequences with diverse sub-events take longer to process and are judged longer than less diverse descriptions or remembered events. This suggests that we reconstruct event duration from memory by retrieving the information associated with the events, providing a sense of the time passed during the actual event unfolding. Therefore, we do not necessarily mentally replay an actual event in its entire temporal development to estimate its duration, but rather, we retrieve structured and schematic memory representations. This thus provides a “temporally compressed” representation, which can in effect be much shorter in processing terms than the actual duration of the event.

It can be argued, however, that the conception of similarity in language and visual encoding is of a different nature. Indeed, in our visual study, similarity structure does not pertain to the multiple situations in which these events can occur over time, as in semantic memory and language, but rather to those event and sub-event properties that appear salient and help rapid encoding. These conceptions of similarity are different because they involve different types of memory representations, those gleaned from few exposures and those built over multiple experiences. Nevertheless, the dissimilar condition in the present learning study can be conceptualized in terms of contextual changes, as in previous cognitive accounts of retrospective temporal estimation (Block, 1992). The sub-events in the dissimilar animations are

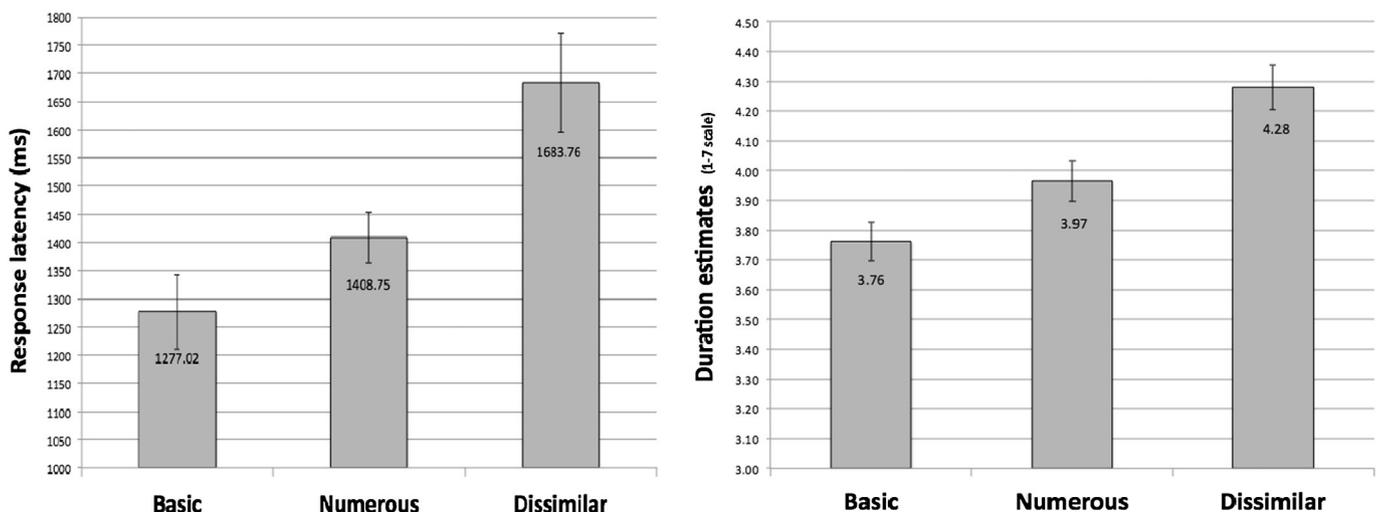


Fig. 2. Response latencies (ms) in the memory task and duration ratings (1–7) in the duration judgment task (error bars indicate standard errors).

not only different from each other but their respective contexts of occurrence are also different (the preceding and subsequent events). Contextual diversity, albeit restricted to few occurrences, is also processed in these cases.

In any case, even if the conception of similarity in language and episodic representations are kept distinct, both cases rely on similarity-based structural analyses and similarity-based encoded representations of experiences. In the language case, events we talk about and regularly observe or perform have been stored in memory over the course of our lives. The binding of event properties into memory schemas depends on recurrent patterns of co-variation between these properties in our experience, which determine the perceived similarity structure of events in the world and are known to lead to schema extraction in cognitive development (Fiser & Aslin, 2005; Orbán, Fiser, Aslin, & Lengyel, 2008; Sloutsky, 2003). In our episodic learning study, the learning task and the minimal nature of the exposure constrained participants to focus on those aspects of the events that serve the current purpose. Yet, similarity structure (chunking together similar occurrences) appears to aid information processing, in the same way that similarity between objects helps categorization and knowledge organization in memory (Goldstone, 1994; Rosch & Mervis, 1975).

Essentially, we argue that these parallels across cognitive domains are not coincidental, but might reveal general similarity-based information processing mechanisms. Event representations emerge from the structural analysis that we apply during encoding and the schema formation (or memory consolidation) processes taking place over time with subsequent exposures. This is not entirely surprising, as our mind must surely impose constraints on the structure of our experience. However, much more needs to be done to establish differences between event representations in episodic and semantic memory and the relationship between the two. Representations of events in semantic memory for example are surely richer and associated with multiple features, from their typical agents and patients, typical locations, etc. (Ferretti, McRae, & Hatherell, 2001; Hare, Jones, Thomson, Kelly, & McRae, 2009), as determined by co-occurrences in previous experience. Episodic representations, in contrast, are specific to one event and may or may not be later integrated with other events in memory. Likewise, more research is needed to establish what exactly counts as similar in different situations, experimental or otherwise. The specific similarity structure of our experiences may depend on different characteristics depending on task demands and the goal of the perceiver, leading to different structural analyses of the information. Similarity for example may depend on the saliency of a difference in a given context, e.g., in Fig. 1, a square disappearing as opposed to moving away, on repeating patterns across different contexts (e.g., typical components of cooking events), or, the dissimilarity of contexts across instances of an element perceived as *the same* or recurrent, as in statistical learning. These instantiations of similarity are different, but all rely on detecting contrasts between upcoming information and that which is already maintained in short or long-term memory, in other words, between bottom-up and top-down information. It is therefore possible, although by no means conclusively demonstrated, that some general similarity-based processing mechanisms may operate across different tasks and contexts.

In sum, we propose that the similarity structure of an event, and in particular, the diversity or dissimilarity between and/or within its associated memory representations plays a critical role in encoding events in memory, in consolidating representations in semantic memory and in the subsequent retrieval cost of these representations in memory or duration estimation tasks. The role of similarity therefore appears as a general mechanism of learning, irrespective of whether similarity is extracted from repeated or limited exposures or from features external or internal to the events. Thus, in retrieving event representations from semantic or episodic memory, the amount of information retrieved, which varies with the similarity of the memory associations involved,

modulates duration estimates. Event properties such as this explain why the actual temporal development of an event maintains some analogical resemblance to our mental representation of it, although what is conceived as duration is not necessarily the same in both the mental and the real worlds.

References

- Belezza, F.S., & Young, D.R. (1989). Chunking of repeated events in memory. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 15(5), 990–997.
- Block, R.A. (1974). Memory and the experience of duration in retrospect. *Memory & Cognition*, 2(1), 153–160.
- Block, R.A. (1982). Temporal judgments and contextual change. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 8(6), 530–544.
- Block, R.A. (1992). Prospective and retrospective duration judgment: The role of information processing and memory. In F. Macar, V. Pouthas, & W.J. Friedman (Eds.), *Time, Action and cognition*. Netherlands: Kluwer Academic Publishers, http://dx.doi.org/10.1007/978-94-017-3536-0_16.
- 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.
- Block, R.A., & Zakay, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, 4(2), 184–197.
- Boltz, M.G. (2005). Duration judgments of naturalistic events in the auditory and visual modalities. *Perception & Psychophysics*, 67(8), 1362–1375.
- 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, <http://dx.doi.org/10.1037/a0016797>.
- Coll-Florit, M., & Gennari, S.P. (2011). Time in language: Event duration in language comprehension. *Cognitive Psychology*, 62(1), 41–79.
- Ferretti, T.R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role. *Journal of Memory & Language*, 44(4), 516–547.
- Fiser, J., & Aslin, R.N. (2005). Encoding multiple element scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology, General*, 134(4), 521–537.
- Gennari, S.P. (2004). Temporal references and temporal relations in sentence comprehension. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 30(4), 877–890, <http://dx.doi.org/10.1037/0278-7393.30.4.877>.
- Goldstone, R.L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, 52(2), 125–157 (Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7924201>).
- Gómez, R., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, 4(5), 178–186.
- 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.
- Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, 111(2), 151–167, <http://dx.doi.org/10.1016/j.cognition.2009.01.009>.
- Joergensen, G., & Gennari, S.P. (2013). Exploring representations of event duration in language. *26th Annual CUNY conference on Human Sentence Processing*.
- Kelter, S., Kaup, B., & Claus, B. (2004). Representing a described sequence of events: A dynamic view of narrative comprehension. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 30(2), 451–464, <http://dx.doi.org/10.1037/0278-7393.30.2.451>.
- Melton, A.W. (1967). Repetition and retrieval from memory. *Science*, 158(3800), 532.
- Metusalem, R., Kutas, M., Urbach, T.P., Hare, M., McRae, K., & Elman, J.L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of Memory and Language*, 66(4), 545–567, <http://dx.doi.org/10.1016/j.jml.2012.01.001>.
- Orbán, G., Fiser, J., Aslin, R.N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences of the United States of America*, 105(7), 2745–2750.
- Ornstein, R.E. (1969). *On the experience of time*. Harmondsworth, England: Penguin.
- Rosch, E., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.
- Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52, <http://dx.doi.org/10.1016/S0010-0277%2898%2900075-4>.
- Sloutsky, V.M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences*, 7(6), 246–251.
- Yonelinas, A.P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 356(1413), 1363–1374.
- 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.
- Zwaan, R.A. (1996). Processing narrative times shifts. *Journal of Experimental Psychology, Learning, Memory, and Cognition*, 22, 1196–1207.
- Zwaan, R.A., & Radvansky, G.A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162–182.
- Zwaan, R.A., Radvansky, G.A., Hilliard, A.E., & Curriel, J.M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, 2(3), 199–220.