Journal of Experimental Psychology: Learning, Memory, and Cognition

Effects of Learned Episodic Event Structure on Prospective Duration Judgments

Myrthe Faber and Silvia P. Gennari

Online First Publication, January 12, 2017. http://dx.doi.org/10.1037/xlm0000378

CITATION
Faber, M., & Gennari, S. P. (2017, January 12). Effects of Learned Episodic Event Structure on Prospective Duration Judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. http://dx.doi.org/10.1037/xlm0000378

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/17/$12.00 http://dx.doi.org/10.1037/xlm0000378

Effects of Learned Episodic Event Structure on Prospective Duration Judgments

Myrthe Faber and Silvia P. Gennari
University of York

The field of psychology of time has typically distinguished between prospective timing and retrospective duration estimation: in prospective timing, participants attend to and encode time, whereas in retrospective estimation, estimates are based on the memory of what happened. Prior research on prospective timing has primarily focused on attentional mechanisms to explain timing behavior, but it remains unclear the extent to which memory processes may also play a role. The present studies investigate this issue, and specifically, the role of newly learned encoded event structure. Two structural properties of dynamic event sequences were examined, which are known to modulate retrospective duration estimates: the perceived number of segments and the similarity between them. We found that when duration and episodic event content are both attended to and encoded, more segments and less similarity between them led to longer attributed durations, despite clock duration remaining constant. In contrast, when only duration is attended to, only the number of segments influenced estimated durations. These findings indicate that incidentally or intentionally encoded episodic event structure modulates prospective duration judgments. Based on these and previous findings, implications for the role of memory mechanisms on prospective paradigms are discussed.

Keywords: episodic memory, event perception, prospective duration estimation, time

Although we do not always have access to a veridical measure of time in terms of seconds or minutes, we are able to judge the duration of an experience in various ways. The psychology of time literature has traditionally distinguished between two main paradigms to study duration judgments. In prospective paradigms, participants are instructed to attend to time during stimulus processing and thus have access to temporal information when judging stimulus duration. In contrast, in retrospective paradigms, participants are not instructed to keep track of time and are unaware of an upcoming temporal judgment, and can therefore only reconstruct duration based on their memory of what happened during a given interval. These paradigms have been argued to involve different cognitive processes, whereby prospective timing primarily depends on attentional mechanisms and retrospective judgments depend on memory processes (Block & Gruber, 2014; Block, Hancock, & Zakay, 2010; Block & Zakay, 1997; Brown, 2008; Grondin, 2010; Grondin, 2001; Zakay, Tsal, Moses, & Shahar, 1994). Prospective judgments in particular have been explained as a process of dynamic attending (Jones & Boltz, 1989) or as a processes of pulse accumulation from an internal clock for a subsequent—often comparative—judgment (Grondin, 2005, 2010; Wearden & Lejeune, 2008). Although internal clock models include a working memory component to allow for comparisons between intervals, much of the prospective research has been devoted to understanding the effects of cognitive load and related variables on timing. When attention is diverted to an additional concurrent task other than timing, duration judgments will vary as a function of the attention devoted to the timing task (Block et al., 2010; Brown, 2008; Brown & Boltz, 2002; Zakay & Block, 1996).

However, recent results have challenged the claim that memory processes other than working memory play a minor or no role in prospective judgments. Indeed, it has been shown that even when we have access to (veridical) clock information, top-down knowledge about duration plays a role in prospective timing. For example, the number of consecutive pop songs played in the background while participants performed a lexical-decision task modulated prospective duration estimates: more remembered songs led to longer estimates, both when people did and did not have access to clock duration (Waldum & Sahakyan, 2013). Similarly, other studies have shown that violations of prior expectations, which require some form of stored knowledge in semantic memory to make predictions, also modulate prospective judgments (Boltz, 2005). Interestingly, in contrast to the number-of-songs effect above, some studies have shown that more segments in the stimuli led to *shorter* judgments, rather than longer, albeit under different task demands (Liverence & Scholl, 2012; Meyerhoff et al., 2015). In any case, it appears that memory mechanisms—and in particular, the influence of prior knowledge on the episodic encoding of an individual experience—play a role in prospective judgments, although the nature of these mechanisms and their relation to task demands remain unclear.

The goal of the present studies was to investigate the role of encoded event structure in prospective duration judgments, and at the same time, shed some light on the effects of different task demands. Unlike previous studies in which the typical duration of songs or prior knowledge of familiar events provided relevant

Myrthe Faber and Silvia P. Gennari, Department of Psychology, University of York.

Correspondence concerning this article should be addressed to Myrthe Faber, who is now at the Department of Psychology, University of Notre Dame, Notre Dame, IN 46556. E-mail: mfaber@nd.edu

information for the prospective time task, we investigated *unfamiliar* dynamic events containing geometric figures moving about or interacting with each other and their environment. The distinction between familiar and unfamiliar stimuli is relevant because stimulus encoding and recollection in these cases may differ considerably and therefore, duration estimations may also differ (Bartlett, 1995; Craik & Lockhart, 1972; Poppenk, Köhler, & Moscovitch, 2010). Indeed, predictable familiar events are encoded and recalled more easily, and duration estimates are shorter compared with unpredictable, hard-to-encode stimuli (Boltz, 2005). Familiar events also have associated typical durations in semantic memory (Coll-Florit & Gennari, 2011), which may influence the duration judgment of a specific event instantiation (Burt & Kemp, 1991). In contrast, novel events for which we have no prior knowledge strongly rely on bottom-up spontaneous structuring to build an episodic representation of the experienced events (Zacks, 2004; Zacks, Speer, Swallow, Braver, & Reynolds, 2007; Zacks & Tversky, 2001). Such newly formed episodic event representations may also modulate prospective judgments, although it is an open question how they do so, particularly when the task—as those used here—requires duration judgments *after* an entire stimulus set has been processed (Boltz, 2005).

The present prospective studies therefore examined the role of newly formed episodic representations of events, the encoding of which is less reliant on familiarity or prior knowledge. In particular, the studies investigated two structural characteristics of events that are known to modulate retrospective duration judgments: the number of perceived segments or event units (whether a dynamic stimulus is perceived as consisting of many or few segments), and the similarity between these segments (Faber & Gennari, 2015a, 2015b). In retrospective paradigms, the similarity structure of the segments matters for event perception because it affects their hierarchical organization into chunks or higher-level segments containing similar lower-level segments, thus influencing the way in which information is encoded. Indeed, repetitive segments can be encoded into one schematic event (Bellezza & Young, 1989; Brady, Konkle, & Alvarez, 2009; Zacks et al., 2007), whereas distinctive segments maintain their status as separate units. Following event structure approaches to memory and perception (Zacks et al., 2007; Zacks, Tversky, & Iyer, 2001), it has been argued that these event structure properties modulate retrospective judgments because they are used to process the stimulus events in the first place and thus they modulate how events are stored and consequently retrieved during duration judgments. The more information is stored and recalled, the longer the attributed duration, as the reconstructed duration is based on the recalled content (Faber & Gennari, 2015a, 2015b; Ornstein, 1969; Poynter, 1983). Thus, retrospective duration judgments of an interval tend to be longer when more segments, rather than fewer, are recollected to have occurred in the interval. Likewise, similar, repetitive segments in an interval can be grouped or chunked into high-level schematic representations during encoding, and thus less information is encoded—and duration judgments tend to be shorter—than when distinctive or individuated segments are separately encoded. In accordance with these encoding structuring principles, Faber and Gennari (2015a, 2015b) found that retrospective duration judgments increased as the number of perceived segments increased and the similarity between segments decreased (i.e., events with numerous and dissimilar segments tend to be judged as longer than those with fewer and/or more similar segments). Given that a retrospective paradigm has already shown effects of memory encoding and retrieval on duration judgments, can such memory modulations also be observed in prospective judgments?

Following these previous studies, the present studies manipulated the event structure of stimulus animations by altering the causal structure of event sequences of the same clock duration. The stimulus animations were arranged into conditions so that for the same clock duration, the conditions varied in the number of perceived segments and their similarity, as determined by independent viewers (see stimulus pretests below). As in prospective paradigms, participants were instructed to attend to the duration of the animations for a later duration judgment task. In Experiment 1, participants were additionally instructed to attend to event content (what happened in the animations) for a subsequent memory test. In Experiment 2, participants were not instructed to study event content and only attended to duration, thus content was encoded incidentally (if at all). To keep the experimental protocol constant, both studies used a repeated-exposure paradigm in which participants associated a stimulus animation with a still frame that was later used as a retrieval cue for the animation. We predicted that if episodic event structure resulting from spontaneous segmentation and perceived similarity during encoding modulates prospective duration judgments, segmental and similarity structure should influence duration judgments to some extent, and possibly, as much as they do in retrospective judgments. The studies therefore were aimed at elucidating the role of encoded event structure in prospective judgments, within the context of the present task demands.

Experiment 1: Intentionally Encoded Event Structure in Prospective Judgments

Experiment 1 examined the role of encoded event structure in prospective duration judgments when participants were instructed to attend to both stimulus duration and stimulus content for subsequent memory and duration tasks. Event structure was manipulated by creating animation triads of the same clock duration that varied in the number of segments and their similarity. The triad members—each representing one experimental condition—displayed the same main geometrical figure moving along a path, but differed in the type of changes undergone by other figures within the animation (e.g., color, shape, path or movement changes; see Figure 1 and examples in https://sites.google.com/a/york.ac.uk/example-animations/). The number of segments perceived in the animation and their similarity were determined in pretest studies (below). According to these judgments, the *basic* condition within the triads contained fewer event segments than the *numerous* condition, whereas the *dissimilar* condition had as many segments as the numerous condition, but less similar ones. Following previous findings showing modulations of encoded structure in retrospective judgments, we expected that similar effects should be found in the present prospective study, given that participants were required to encode the animation content as well as to attend to duration. In particular, if the effect of encoded structure resembles that found for retrospective judgments, we expected a positive trend across conditions so that as the number of segments increased, and the degree of similarity decreased, longer duration judgments should be observed (Faber & Gennari, 2015a).

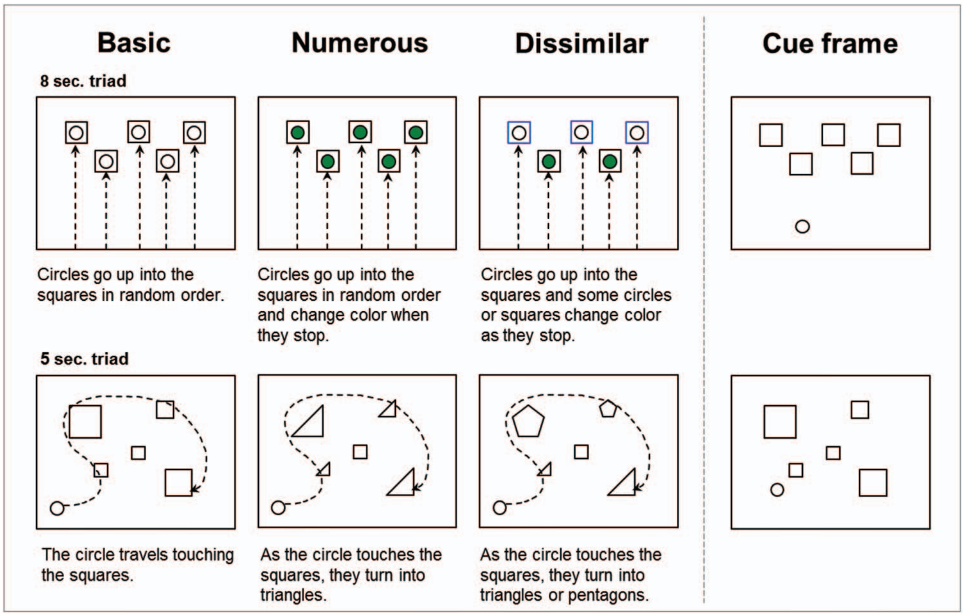

Figure 1. Examples of stimulus triads and cue frames in Experiment 1 and 2. See the online article for the color version of this figure.

Method

Participants. Sample size was based on our previous retrospective study that used the same stimuli and conditions (Faber & Gennari, 2015a). This study yielded a medium effect size (η^2 = .07, observed power: .85) using a sample of 75 participants. Therefore, in this study, we aimed for a similar sample size (25 participants for each of three lists, see below), after excluding participants with poor memory accuracy. A total of 82 native English-speaking students from the University of York (74% female, mean age = 21.2, with normal or corrected-to-normal vision) participated for course credit, course requirement or a small monetary reward. Seven participants with poor memory accuracy were excluded as they had low accuracy scores in the recognition task, according to a criterion previously used in the retrospective studies (recognition accuracy ≤50% in one of the conditions or a false alarm rate above 50%; Faber & Gennari, 2015a). These exclusions were motivated by our goal of examining performance on duration judgments when participants have performed reasonably well in a memory task. If they had no recollection of the animations' content, it is impossible to draw conclusions about the role of encoded event structure. This experiment was approved by the Ethics Committee of the Department of Psychology of the University of York. Participants provided informed consent and were debriefed after the study.

Stimuli. The stimuli were taken from Faber and Gennari (2015a) and consisted of 28 animation triads that were created with Adobe Flash CS5.5, each lasting an integer number of seconds (varying between 3 and 9 seconds, 4 animation triads for each of the 7 time bins). Triad members were equal in clock duration, and were arranged into three conditions: a basic event sequence containing a repeating or stable motion of a shape (*basic condition*), a sequence with a repeating change (e.g., displacement) added onto the basic motion (*numerous condition*), and a sequence like the numerous one, but with dissimilar changes (e.g., displacement and disappearance; *dissimilar condition*; see Figure 1 and examples in https://sites.google.com/a/york.ac.uk/example-animations/). The basic condition in a triad was systematically modified into the numerous condition, which in turn was modified into the dissimilar condition, keeping speed of motion constant. Across triads, shapes, motion, and changes were visually different to prevent memory interference. The stimuli also included a single still frame from near the beginning of the triad animations that was common to all triad members. These cue frames were used as a retrieval cue in the memory and duration judgment task. Two additional anchor animations (lasting 2 and 10 seconds respectively) and their corresponding cue frames were also created to facilitate the duration judgment task, providing the boundaries of a duration scale from the shortest to the longest studied animation (see Design and procedure).

Stimulus pretests. To ascertain that the created animations differed in their stimulus properties across conditions, separate sets of participants provided judgments on the perceived number of segments and the relative similarity between subevents in an animation. Two Web-based questionnaires were conducted. Stimuli were organized in three lists as in the main study (see Design and procedure). A total of 121 English speakers recruited through Amazon Mechanical Turk completed the questionnaires (4 participants were excluded from the segmentation data due to their idiosyncratic scores). A total of 87 participants were used in the segmentation task (29 per list, mean age = 34.6, 52% males) and a total of 30 participants were used in the similarity task (10 per list, mean age = 38.7, 46% males). The web-link provided to participants directed them to a custom-built web page containing a list of stimulus animations.

In the *subevent* questionnaire, participants indicated the number of instances in which a smallest natural and meaningful unit within the animation finished and another started. These instructions were based on those used in event segmentation studies (Zacks et al., 2001). Participants were told to watch the animation several times and count these instances. As shown in Table 1, larger numbers of subevents were perceived in the numerous and dissimilar conditions compared with the basic condition. Repeated measures ANOVA with triads as a random factor and mean number of subevents as dependent variable indicated a main effect of condition ($F(2, 54) = 5.53$, $p = .007$; $\eta^2 = .17$) with all pairwise comparisons being significant (all ps $< .05$) except for that between the numerous and the dissimilar conditions, as expected. Note, however, that the difference in number of segments is relatively small compared with other segmentation studies, in part because large differences between animations of different clock duration (say, 3 vs. 9 seconds) are averaged out in Table 1, as indicated by the large ranges. This motivates the by-item hierarchical regression approach (see Analyses), which better accounts for differences in the number of segments in a given animation.

The *similarity* 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*). Examples were provided illustrating the extreme points of the scale. The animations could be watched as many times as desired by clicking a play button. The order of the animations in the web page was random. Table 1 shows the similarity ratings for each condition. Animations in the numerous condition were judged to contain similar subevents to a comparable extent as the subevents of the basic condition. In contrast, animations in the dissimilar condition were judged to contain less similar subevents. Repeated measures ANOVA with triads as a random factor and similarity rating as dependent variable indicated that there was a main effect of condition ($F(2, 54) = 27.05$, $p < .001$; $\eta^2 = .50$), and all pairwise comparisons were highly significant (all ps $< .01$) except for that between the basic and the numerous conditions, as expected.

Design and procedure. Items were arranged in three lists of 28 animations. Each list contained only one member of each triad (either basic, numerous or dissimilar) but all conditions across triads. The allocation of triad members to lists was as follows. The triads were numbered from 1 to 28 and organized according to duration, with the first four triads lasting 3 sec, the next four triads lasting 4 sec, and so on. For the first triad, the basic, numerous, and dissimilar conditions were assigned to lists 1, 2, and 3, respectively; for the second triad, these conditions were assigned to lists 2, 3, and 1, respectively; for the third triad, they were assigned to lists 3, 2, and 1, respectively and so on. This meant that each list contained either 9 or 10 members of each condition spanning the whole range of clock durations (from 3 to 9 sec). This arrangement was the same as in previous retrospective studies. Participants were then randomly assigned to one of the three lists.

Participants were recruited to take part in an experiment on time and memory, and were instructed to study the content of the animations (i.e., what happens in the animation) and to remember their relative duration, together with their paired frames, which were to be used in a subsequent memory test to probe their memory for the animation content and duration. They were instructed to attend to the relative duration of animations rather than their real-time duration, and were instructed not to tap or count in their head as a way of timekeeping while studying the animations.

In the *study task*, a list of 28 animations plus the two anchor animations (30 animations in total) was presented. Each participant saw their list three times in different random orders. For each trial, the cue frame was displayed for 2 seconds followed by the associated animation. When the animation finished, a new display prompted for a key press to progress to the next trial. The study phase lasted approximately 15 min.

After the study phase, participants performed the *recognition memory task*. The 30 cue-frames that participants had studied plus 30 foil frames (similar frames to those studied but previously unseen) were presented in random order. The foil frames were extracted from a triad member not studied in the list (e.g., if the basic condition was studied for a triad, the foil frame was a middle or late frame from the numerous or dissimilar conditions of that triad), thus keeping foils and cue frames fairly similar and requiring detailed knowledge of the animations. Participants indicated whether each frame belonged to one of the studied animations by pressing a YES or NO key. The number of cue frames and foil frames was equal. A fixation cross was presented on the screen for 1–3 seconds to randomly vary intertrial time between participants' responses. The recognition memory task lasted approximately 5–6 min.

Finally, in the *duration judgment task*, participants were shown the studied cue frames in random order and were asked to indicate the relative duration of their associated animations on a 1–7 scale (1 = *very short*, 7 = *very long*). Instructions indicated (by displaying the anchor animations' cue frames) that the anchor animations were the shortest and longest in the studied set and were outside the scale. The objective was to indicate that the whole 1–7 scale should be used in providing the judgments. Randomly varied intertrial times were used as in the recognition task. In each trial, participants saw the cue frame in the center of the screen, with the 1–7 scale displayed at the bottom. Extreme of the scale were indicated by 1 - *very short*, and 7 - *very long*. This screen was displayed until a participant pressed a number between 1 and 7. A central cross was displayed during the variable intertrial time. Note that alternative tasks used in timing studies were deemed less appropriate for use here because clock estimates (e.g., seconds) encourage counting during stimulus processing, even when participants are told not to do so. Here participants were explicitly told

Table 1
Stimulus Properties

	Number of segments			Segmental similarity		
Conditions	Mean	*SD*	Range	Mean	*SD*	Range
Basic	4.49	1.79	1.79–8.76	5.06	1.26	2–7
Numerous	4.73	1.87	2.14–9.07	5.07	.94	3–6
Dissimilar	4.85	1.85	2.28–9.31	3.89	1.14	2–6

that they would later have to compare duration within the stimulus set and clock-type values would not be requested, and thus, there was no need to count seconds. Moreover, if participants were to count and occupy their processing resources in such a task, there may be interference between counting and encoding content into episodic memory, an issue that we wanted to avoid. We will come back to various possible effects of task demands in the general discussion.

Data treatment.

Recognition memory task. From this task, we only analyzed response latencies to the probes (not foils). Outlier recognition latencies longer than 2.5 standard deviations from each condition mean were excluded from these data (63 of 2100 response latencies, 3% of the data). We excluded one additional response because it was deemed unreliable (<100 ms). For the analyses across conditions, only correct responses were taken into account (i.e., yes-responses to probes).

Duration judgment task. We excluded responses longer than 10 seconds (20 out of 2100 responses, 0.9% of the data), as these were not deemed to represent confident judgments. From these data, 35 responses (1.7%) were also excluded if a participant failed to both correctly recognize and reject the frames associated with an item animation (i.e., an item's cue frame and foil frame were incorrectly rejected and accepted, respectively), as this suggested that there was no appropriate memory trace for that animation.

Analyses. To statistically examine the effect of event structure, we performed two sets of analyses. First, we investigated whether duration ratings across subjects (F_1) and triads (F_2) display the same increasing pattern across stimulus conditions (i.e., manipulations of internal event structure) that we found previously for retrospective estimates. In particular, after the main analysis of variance across conditions, we specifically tested that positive trends across conditions indeed explained most of the variance in the data, as predicted. For this, we conducted planned contrasts for positive trends, which assign the weights -1, 0, and $+1$ to the basic, numerous, and dissimilar conditions, respectively, and then checked that the residual variance of this trend was not significant (Hays, 1981). Analyses performed on raw mean estimates, as well as on mean ratios (estimated duration divided by actual duration) yielded a similar pattern of results (Boltz, 1995). For this latter analysis, the numbers on the 1–7 scale were transformed into clock times as indicated by the duration of the anchor animations (e.g., 1 representing 3 seconds, up to 7 representing 9 seconds) and then divided by the actual duration of the animation.

Second, we used hierarchical multiple regression analyses to assess the independent contribution of clock duration and event properties on an animation-by-animation basis, which captured the effect of the items' individual event properties more precisely than grouping conditions. These analyses allowed examining the proportion of variance accounted for at each step of the model over and above other predictors by computing R^2 change statistics. For instance, if similarity between subevents explains variance independent of segmentation, it should significantly increase the variance accounted for by the previously entered predictors, such as the number of segments and clock duration.

Results

Recognition memory task. The overall recognition memory accuracy was 87% (SD = 8.8%) after participant exclusions, taking correct identification into account (see Table 2). There was no difference in hits between conditions (Friedman's test $\chi^2(2)$ = .032, *ns*). Note that accuracy of correct recognition in this task was no different from that in our previous retrospective study (Mann–Whitney test, U = 2589, *ns*), suggesting that the instruction to attend to duration did not preclude participants from encoding the content of the animations as well as in a retrospective paradigm.

The recognition time data replicated previous findings, with recognition times increasing as a function of condition (see Table 2). The results of repeated measures ANOVAs with either subjects (F_1) or triads (F_2) as a random factor and recognition times as a dependent variable indicated that there was a significant main effect of condition on the response latencies, $F_1(1.78, 132^*) = 3.40$, $p = .041$, $\eta^2 = .044$ *Greenhouse-Geisser corrected for sphericity; $F_2(2, 54) = 3.49$, $p = .037$, $\eta^2 = .115$, and significant positive trends, $F_1(1, 74) = 4.80$, $p = .032$, $\eta^2 = .061$, $F_2(1, 27) = 8.33$, $p = .008$, $\eta^2 = .236$. These differences reflect the role of qualitative aspects of the animations (Yonelinas, 2001), suggesting that properties of the events encoded in memory during the study phase were accessed or checked during memory judgments, with the conditions that required access to more segments and less similarity between them leading to increased response latencies. Because the implications of these recognition findings have been discussed in detail elsewhere, we do not discuss them further here (Faber & Gennari, 2015a).

Duration judgment task. We expected that if encoded event structure modulated prospective duration judgments, the pattern of results should resemble that found in the parallel retrospective study previously reported: duration ratings should vary across conditions following a positive trend, despite clock duration remaining constant. Mean ratios across conditions differed significantly, $F_1(2, 148) = 7.82$, $p = .001$, $\eta^2 = .10$; $F_2(2, 54) = 4.26$, $p < .02$, $\eta^2 = .14$[1] (see Table 3). The dissimilar condition was judged longer than the other conditions, and the numerous condition was judged longer than the basic condition. As predicted, the duration ratios displayed a positive trend, $F_1(1, 74) = 11.96$, $p = .001$, $\eta^2 = .14$; $F_2(1, 27) = 6.67$, $p = .02$, $\eta^2 = .20$, with no significant residual variance in the data, $F_1(1, 74) = 1.38$, $p = .76$; $F_2(1, 27) = .33$, $p = .43$. Analyses using the raw duration ratings yielded a similar main effect of condition, $F_1(2, 148) = 3.994$, $p = .02$, $\eta^2 = .051$ across subjects; $F_2(2, 52) = 3.673$, $p = .03$, $\eta^2 = .12$, controlling for differences across clock durations by adding clock duration to the model as a covariate, and displayed a similar positive trend, $F_1(1, 74) = 7.26$, $p = .009$, $\eta^2 = .089$; $F_2(1, 26) = 7.31$, $p = .01$, $\eta^2 = .22$. These findings suggest that in the present prospective paradigm, longer durations were attributed to events of the same clock duration as a function of event structure.

Regression analyses. To evaluate the effect of segmental and similarity characteristics of the animations on duration judgments, we conducted by-item hierarchical multiple regressions and examined the proportion of variance accounted for by segment and

[1] Note that a similar ANOVA conducted with experimental list as a factor revealed no main effect of list or interaction with condition (list effect: $F_1(2,72) = 1.22$, $p = .30$), interaction: $F_1(4,144) = .26$, $p = .90$).

Table 2
Recognition Performance and Recognition Times for Experiment 1

Measure	Condition		
	Basic	Numerous	Dissimilar
Correct recognition	88%	88%	87%
Recognition times (ms)	1291 (45)	1367 (54)	1400 (57)

Note. Standard errors in parentheses.

Table 4
Multiple Regression Model for the Duration Ratings of Experiment 1

Model	*B*	*SE B*	β
Model 1			
Constant	2.54	.23	
Clock duration	.27	.04	.63***
Model 2			
Constant	2.32	.22	
Clock duration	.15	.05	.34**
Number of sub-events	.21	.05	.43***
Model 3			
Constant	2.93	.35	
Clock duration	.14	.04	.32**
Number of sub-events	.20	.05	.43***
Similarity	−.12	.05	−.17*

* $p < .05$. ** $p < .01$. *** $p < .001$.

similarity scores (obtained from the independent pretest studies). Note that there was no significant correlation between the number of segments and similarity scores, $r = -.09$, $p = .40$. In the first step of the regression model, we included clock duration as a control predictor to account for the systematic variation built across triads. This regression model thus contained mean ratings per item as a dependent variable and clock duration as predictor. In the second step, we added subevent scores to the first regression model. This yielded a significant increase in the proportion of variance accounted for: R^2 increased from .39 to .50, $F_{change}(1, 81) = 16.86$, $p < .001$. In the third step, we added the similarity ratings to this regression model, which also yielded a significant increase in the proportion of variance accounted for: R^2 increased from .50 to .53, $F_{change}(1, 80) = 4.75$, $p = .03$. This pattern of significance remained the same regardless of the order in which the predictors were added to the model. Table 3 reports the statistics of the full regression model. Thus, although as expected, clock duration was able to explain a significant proportion of variance in the duration ratings, both the number of perceived subevents and similarity between them significantly improved the fit, suggesting that participants attributed duration based on both attended time and the event structure encoded in memory: increasing the number of subevents in a sequence led to increased duration attributions, and so did decreasing the similarity between them (hence, the negative relationship in Table 4).

Decision times. For completeness, we also examined the time it took participants to provide a duration rating across conditions, although we did not predict an effect of condition in this measure because the parallel retrospective study in Faber and Gennari (2015a) did not yield any differences in decision times. Note also that our decision times reflect decision-making processes leading to a rating—in particular, comparisons across previously encoded events relative to a scale—and little is known about how event segmental structure may modulate such decisions. Unlike duration reproduction and perceptual judgments immediately after a stimulus is perceived, rating decisions in this experiment must rely on previously encoded memory cues, and might involve currently unexplored processes, such as competition between alternative memory cues or potential ratings, or even heuristic strategies (Oppenheimer & Kelso, 2015; Shah & Oppenheimer, 2008).

The decision times for the ratings reported above were used in the analysis. Results indicated that there was no main effect of condition on decision times (mean decision times: basic condition = 3051 ms, numerous condition = 3004 ms, dissimilar condition = 3054 ms; $F_1(1.84, 136^*) = .237$, $p = .771$; $F_2(1.65, 44.6^*) = .277$, $p = .717$ *Greenhouse-Geisser corrected for sphericity). Furthermore, there was no significant relationship between the decision times and duration ratings on a by-animation level ($r = .012$, $p = .91$; $r_{partial} = -.021$, $p = .85$, controlling for clock duration). The absence of a condition effect was further corroborated by exploratory by-animation regression analyses similar to those above, which revealed no relationship between mean decision times and actual clock duration or number of segments, suggesting that decision times are not modulated by those variables that more strongly predicted attributed duration. We did however observe a significant negative relationship between similarity and decision times ($\beta = -.242$, $p = .029$). which might be attributable to alternative ratings or memory cues being considered in making decisions for animations with less similar segments. These results indicate that attributed durations (i.e., the ratings) and decision times were driven by different animation properties and reflect different aspects of processing. Most notably, decisions times were unrelated to actual clock duration, a typical property of duration judgments. Therefore, we focus our discussion on the duration rating ultimately attributed to an animation, rather than decision times, which appear to reflect decision processes.

Discussion

The results of Experiment 1 demonstrated that prospective duration judgments—but not necessarily decision times—were modulated by event properties when participants were required to remember both relative duration and event content. Both the number of perceived segments and the similarity between them explained significant proportions of variance in the duration ratings, over and above clock duration and over and above each other. An increase in the number of segments and a decrease in similarity were related to an increase in duration rating, as expected. Moreover, the pattern of performance in the memory recognition task

Table 3
Duration Ratings Across Items for Experiment 1

Measure	Condition		
	Basic	Numerous	Dissimilar
Duration ratings (1–7 score)	4.03 (.18)	4.17 (.16)	4.24 (.16)
Duration rating ratios	1.10 (.06)	1.13 (.07)	1.15 (.07)

Note. Standard errors in parentheses.

was not different from that in a retrospective paradigm, suggesting that content was encoded equally well in both paradigms regardless of additional attention to time in the present experiment. Importantly, the present recognition and judgment results clearly parallel those previously reported using a retrospective paradigm (Faber & Gennari, 2015a), and suggest that remembering episodic event structure affects prospective duration judgments if participants are instructed to encode content. Therefore, prospective judgments *can* be modulated by specific aspects of episodic memory content, suggesting a role for encoding and retrieval mechanisms, at least in some circumstances.

It is interesting to note that attending to time did not appear to have interfered with encoding content, which contrasts with previous suggestions by some attention models (Brown, 2008). Attending to time did not preclude from encoding event content likely because the two tasks were not in conflict, and attention did not need to be alternatively allocated to one or another task goal. In fact, as suggested by some timing models, it is possible that attending to content, for example, the number of segments perceived, helps timing, and can be used as a cue to later recollect and evaluate duration (Jones & Boltz, 1989). This is very similar to what has been proposed for retrospective paradigms (Block, 1982; Ornstein, 1969; Poynter, 1989) and suggests that recollection mechanisms appear to play a role in prospective paradigms too, at least when stimulus content is attended to.

Experiment 2: Incidentally Encoded Event Structure in Prospective Judgments

Experiment 1 indicated that the episodic event structure encoded during stimulus processing played a role in prospective time judgments when attention to stimulus content is also required by the task instructions. However, it remains unclear whether episodic structure would still play a role when attention to content is not required for a subsequent memory task. Therefore, to establish whether episodic event structure also plays a role when no attention to content is required, we conducted an experiment exactly as Experiment 1, but only instructed participants to attend to duration. Importantly, exposure to the stimuli was the same as in Experiment 1, thus making incidental encoding possible. As before, if episodic memory encoding and retrieval play any role in prospective duration judgments, we would expect modulations of event structure on duration judgments, and specifically, of the number of segments and their similarity.

Method

Participants. Participants were 79 native English-speaking students from the University of York (70% female, mean age = 21.0) who participated for course credit, course requirement, or a small monetary reward. Four participants with poor memory accuracy were excluded as they had low accuracy scores in the recognition task (overall recognition accuracy ≤50% or a false alarm rate above 50%), resulting in 75 participants in total (25 per list). As study of the animations' content was not required, we expected recognition to be lower than in Experiment 1. The exclusion criteria used here were therefore slightly more lenient than in Experiment 1 (overall recognition accuracy ≤50% instead of ≤50% in one or more conditions). Participants had normal or corrected-to-normal vision. This experiment was approved by the Ethics Committee of the Department of Psychology of the University of York. Participants provided informed consent and were debriefed after the study.

Stimuli. The same animations were used here as in Experiment 1.

Design and procedure. The design of this study was like that of Experiment 1, except that participants were instructed to pay careful attention only to the relative duration of the animations. As before, they were instructed not to tap or count in their head as a way of time keeping while watching the animations, as judgments in seconds would not be required. Importantly, participants received no instruction to remember the content of the animations. It was pointed out to them that each animation would be preceded by a still frame that would later be used to refer back to the animation. After the study phase, participants performed a recognition memory task, which to these participants came as a surprise. Finally, they performed a duration rating task identical to that of Experiment 1.

Data treatment.

Recognition memory task. For analyses of the recognition memory task, as in Experiment 1, outlier recognition latencies longer than 2.5 standard deviations from the condition mean were excluded from these data (47 out of 2100 response latencies, 2.1% of the data).

Duration judgment task. As in Experiment 1, items for which it took a participant longer than 10 seconds to respond (59 out of 2100, 2.8% of the data) were excluded from the duration judgments. To keep exclusions comparable across experiments, 53 responses (2.5%) were also excluded if a participant failed to both correctly recognize and reject the frames associated with an item animation (i.e., an item's cue frame and foil frame were incorrectly rejected and accepted, respectively). However, given that memory accuracy was lower than Experiment 1 (see below) we also conducted all analyses without this memory exclusion, yielding similar results.

Results

Recognition memory task. The overall recognition memory accuracy was 83% (*SD* = 15%) taking correct identification into account. There was no difference in correct identification between conditions (Friedman's test, $\chi^2(2) = 2.86$, *ns*; see Table 5). However, correct identification was significantly lower in Experiment 2 compared with Experiment 1 (Mann–Whitney test, $U = 2091.50$, $p = .007$), suggesting that event structure was less deeply encoded here compared with carefully studying the animations for a subsequent memory test as in Experiment 1. This was expected given the varying task instructions and the unexpected nature of the recognition task for the present participants.

A similar suggestion was evidenced by the pattern of recognition times, which did not vary across conditions. The results of repeated measures ANOVAs with either subjects (F_1) or triads (F_2) as a random factor and recognition times as a dependent variable indicated that there was no significant main effect of condition on the response latencies, $F_1(2, 148) = .471$, $p = .63$, $\eta^2 = .006$; $F_2(2, 54) = .26$, $p = .78$, $\eta^2 = .009$, no significant positive trends, $F_1(1, 74) = .007$, $p = .94$, $\eta^2 = .000$, $F_2(1, 27) = .061$, $p = .81$, $\eta^2 = .002$, and no significant contrast across

Table 5
Recognition Memory Results for Experiment 2

Measure	Condition		
	Basic	Numerous	Dissimilar
Correct recognition	81%	85%	83%
Recognition times (ms)	1386 (66)	1426 (65)	1382 (61)

Note. Standard error in parentheses.

conditions (all p's $> .05$). Taken together, these results suggest that participants' memory of the content of the animation was poorer compared with Experiment 1, as expected.

Duration judgment task. We expected that if the number of subevents and their dissimilarity modulate the attributed duration of events, the duration ratings should vary across conditions, despite clock duration remaining constant. Table 6 displays the mean ratios across conditions, which did not differ significantly, $F_1(1.66, 122.4^*) = 1.37$, $p = .26$, $\eta^2 = .02$ *Greenhouse-Geisser corrected for sphericity; $F_2(2, 54) = .35$, $p = .71$, $\eta^2 = .01$. There were no significant positive trends, $F_1(1, 74) = 2.59$, $p = .11$, $\eta^2 = .03$; $F_2(1, 27) = .57$, $p = .46$, $\eta^2 = .02$, and no significant contrasts between conditions. This pattern of results remained the same if the raw rating scores were used as dependent variable, with no significant main effects of condition, $F_1(2, 148) = .508$, $p = .60$, $\eta^2 = .007$ across subjects; $F_2(1, 52) = .030$, $p = .97$, $\eta^2 = .001$, across triads controlling for differences across clock durations by adding clock duration to the model as a covariate, and no significant positive trends, $F_1(1, 74) = 1.01$, $p = .319$, $\eta^2 = .013$; $F_2(1, 26) = .004$, $p = .95$, $\eta^2 = .000$.

Note that the exclusion criteria based on memory performance used here were more lenient than those used in Experiment 1. We also performed the above analyses using the same memory exclusion criteria as Experiment 1. This resulted in the exclusion of 11 participants based on memory performance, and another five (with the next lowest memory scores for their respective lists) to counterbalance the number of participants per list (total sample size of 63, resulting in 21 participants per list). Given the effect size of Experiment 1 ($\eta^2 = .10$, observed power $= .95$), a sample size of 41 subjects should be sufficient to detect an effect of condition in a prospective paradigm (Faul, Erdfelder, Lang, & Buchner, 2007). No significant differences were found between the mean rating ratios across conditions, $F_1(1.72, 106.4)^* = 1.36$, $p = .26$, $\eta^2 = .02$ *Greenhouse-Geisser corrected for sphericity; $F_2(2, 54) = .29$, $p = .75$, $\eta^2 = .01$. There were no significant positive trends, $F_1(1, 62) = 2.35$, $p = .13$, $\eta^2 = .04$; $F_2(1, 27) = .26$, $p = .62$, $\eta^2 = .009$, and no significant contrasts between conditions. The same holds true for analyses of the raw duration ratings, displaying no main effect of condition, $F_1(2, 124) = .612$, $p = .54$, $\eta^2 = .01$ across subjects; $F_2(2, 52) = .439$, $p = .65$, $\eta^2 = .017$ across triads controlling for differences across clock durations by adding clock duration to the model as a covariate, and no significant positive trends or contrasts (all $p \geq .32$). Although we acknowledge that the true effect size could be smaller, the present findings suggest that there was no significant effect of condition on duration ratings.

Regression analyses. Because regressions on an item-by-item basis can be more sensitive to establish correspondences across continuous variables than a categorical ANOVA, particularly considering that the animations varied in duration, number of segments and similarity within a given grouping category, we conducted the same hierarchical multiple regression analyses performed in Experiment 1 to test whether event structure modulated duration ratings. In the first step, clock duration was added to the model as a control predictor. The number of segments was then added to this model, yielding a significant increase in the proportion of variance accounted for: R^2 increased from .27 to .37, $F_{change}(1, 81) = 12.75$, $p = .001$. The similarity ratings were then added to this model, but did not significantly increase the proportion of variance explained (R^2 remains .37). Table 7 reports the statistics of the full regression model. Results did not change when adding the variables in any different order, or when entering data with the stricter participant exclusions as a dependent variable. Thus, the number of subevents but not their relative similarity was related to an increase in attributed duration when participants were not required to encode content.

Decision times. For completeness, we examined the time it took participants to provide a duration judgment across conditions, as done for Experiment 1. Although we did not have specific predictions for decision times based on previous findings, it might be informative to compare the results across the two experiments, which only differed in task instructions. Similar to the decision times of Experiment 1, results indicated that there was no significant main effect of condition, $F_1(2, 148) = 1.59$, $p = .21$; $F_2(2, 54) = 2.15$, $p = .13$. The dissimilar condition showed numerically longer decision times (mean basic condition = 3088 ms, numerous condition = 3097 ms, dissimilar condition = 3220 ms). Further by-item regression analyses similar to those above revealed that decisions times were not predicted by clock duration and segmental structure—which did explain duration ratings—but did show a negative relationship with similarity ($\beta = -.222$, $p = .047$). As in Experiment 1, decision times were not related to clock duration and number of segments as attributed duration was, suggesting that decisions times were orthogonal to attributed duration.

Finally, the decision times in Experiment 2 were not substantially different from those of Experiment 1. A mixed ANOVA with mean decision times as dependent variable, condition as repeated factor and experiment as between-participants factor indicated no main effect of condition, $F_1(1.81, 296^*) = 1.22$, $p = .30$; $F_2(1.75,108^*) = 1.85$; $p = .16$ *both Tests Greenhouse-Geisser corrected for sphericity, no main effect of experiment, $F_1(1, 148) = .37$, $p = .54$; $F_2(1, 54) = .58$, $p = .44$, and no interaction. Direct comparisons between experiments for each condition separately indicated no significant difference across experiments (all $ps > .05$). This suggests that task instructions did not significantly alter the overall length or difficulty of the decision process across experiments.

Table 6
Duration Ratings Across Items for Experiment 2

Measure	Condition		
	Basic	Numerous	Dissimilar
Duration ratings (1–7 score)	4.16 (.16)	4.22 (.15)	4.26 (.17)
Duration rating ratios	1.14 (.07)	1.14 (.07)	1.15 (.07)

Note. Standard errors in parentheses.

Table 7
Multiple Regression Model for the Duration Ratings of Experiment 2

Model	*B*	*SE B*	β
Model 1			
Constant	2.93	.25	
Clock duration	.21	.04	.52***
Model 2			
Constant	2.72	.24	
Clock duration	.10	.05	.23§
Number of sub-events	.20	.06	.42***
Model 3			
Constant	2.73	.40	
Clock duration	.10	.05	.23§
Number of sub-events	.20	.06	.42***
Similarity	−.001	.06	−.002

*** $p \leq .001$. § $p \leq .06$ (i.e., marginally significant).

Discussion

The results of Experiment 2 suggest that there was no effect of grouping condition on duration ratings when participants have explicitly attended to relative duration but not to stimulus content. Less attention to the specific animation content was evidenced in relatively poor recognition memory performance, compared with Experiment 1, and the absence of a condition effect in recognition times—an effect replicated several times when content is attended to (Faber & Gennari, 2015a, 2015b). Because our grouping conditions were based on the number of fine-grained—perceived as the smallest—segments and their similarity, which require attention to specific details of the animations (e.g., color or shape changes), the absence of a condition effect on duration ratings is not surprising, as these specific aspects may have not been attended to, encoded, or recollected. However, the regression results indicated that across items, the number of segments (but not the similarity between them) was a good predictor of attributed duration, with higher numbers of segments being related to longer duration ratings, over and above clock duration. Together, these findings suggest that the effect of episodic event structure on duration ratings was reduced but present when content is not explicitly encoded.

Importantly, the observed positive effect of the number of segments on the duration ratings aligns with our prediction that recollection of more segments during duration judgments leads to longer ratings, and is thus consistent with an influence of memory content on prospective judgments, even when participants are not instructed to remember content. Surely repeated exposure to the stimuli led participants to learn something about the animations. But what do they attend to and encode under the present instructions? One possibility is that when attending to time, at least some salient event boundaries are important for the perception of the rhythm or temporal development of the events, and thus are potentially integral to the perception of duration, as previously suggested (Boltz, 1992; Jones & Boltz, 1989). The event boundaries attended to in Experiment 2 are likely to be of a more coarse size than those in Experiment 1 and our pretest segmentation study. What happens at fine-grained event boundaries (e.g., whether a shape changes color/texture, or whether a minor shape moves) is more relevant to a detailed representation of content (as in Experiment 1), but these low-level segments might not be attended to as much in Experiment 2 because fine-grained boundaries are not considered relevant when encoding information for a duration judgment task. These observations are consistent with many event segmentation studies suggesting that fine-grained segmentation leads to better recall of specific event characteristics than coarse segmentation, a consequence of *orienting* or attending to larger or smaller segments in the stimulus stream (Hanson & Hirst, 1989; Zacks et al., 2007, 2001).

A difference in orientation to the stimuli in Experiment 1 and 2 (encoding specific stimulus characteristics or not) therefore explains why we observed relatively poor memory recognition performance and no effect of similarity or grouping conditions in the durations ratings of Experiment 2, but we did observe an effect of number of segments in our regressions: only a coarse segmentation structure was incidentally encoded into memory, possibly as part of the timing process. Because coarse segments tend to include fine-grained segments and both are related to salient changes (e.g., lower-level segments are embedded into higher-level segments), coarse and fine-grained segments are typically correlated (Zacks, Speer, & Reynolds, 2009). Therefore, a correlation between the number of fine-grained segments and duration judgments would nevertheless be expected in Experiment 2, as observed in the regression analyses. We therefore conclude that despite different attention demands across our studies, the results of rated duration in Experiment 2 suggest an influence of coarse stimulus segmentation during encoding, which is then recollected and used in duration judgments.

General Discussion

The findings of Experiment 1 indicated that episodic event structure, as reflected by the number of event segments and their similarity, modulated duration ratings in a prospective paradigm when the content of the interval had been encoded. Moreover, this study showed that the same fine-grained features of event structure that modulated duration ratings in a retrospective paradigm also played a role in a prospective paradigm. This clearly indicates that similar memory processes are involved in both prospective and retrospective judgments when stimulus content is attended to. The findings of Experiment 2 then revealed that the influence of encoded episodic event structure on prospective judgments is reduced when attention to stimulus content is not required, as evidenced by relatively poor memory performance and no effect of grouping conditions on duration judgments. Nevertheless, the number of fine-grained segments in the stimuli—which tend to correlate with larger coarse segments—accounted for a significant proportion of variance in duration judgments over and above actual clock duration. This result suggests a role for segmentation when attending only to duration, despite more superficial stimulus encoding. Therefore, the results overall indicate a role for encoded event structure in prospective duration judgments.

Because effects of stimulus structure are ultimately linked to the way in which the stimuli were segmented and encoded during learning, it can be inferred that episodic memory content modulated the present prospective judgments. When a memory task was expected along with duration judgments, participants attended to fine-grained stimulus properties, including small recognizable units (segments) and their specific characteristics (changes in shape, color, movement, etc.). After learning, at the point of the duration judgment, participants retrieved or automatically activated the animation content and used this information to provide a duration rating. As in retrospective studies, the more information is recollected (more segments and more distinctive properties), the longer the judgments. Critically, when only a (comparative) duration judgment task was expected, the influence of segmental structure remained. In this case, participants may have attended to coarse stimulus properties during exposure—those considered relevant for the duration rating task—and therefore, the encoded segmental structure—but not segmental identity—was retrieved and used to provide a duration judgment: the more segments encoded and recalled, the longer the duration ratings. This suggests that some aspects of the stimulus content, even if coarse, are potentially integral to attending to duration in at least some prospective paradigms, because the stimulus content provides relevant clues to judge duration, as suggested by the dynamic attending model (Jones & Boltz, 1989).

These findings contrast with frequent claims that prospective judgments are not modulated by encoded information or memory recollection (Block & Gruber, 2014; Block et al., 2010; Zakay & Block, 1997), and argue for a more nuanced view of the role of memory in prospective timing: when available, we employ our memory of event structure to inform our duration judgments, even when we have incidentally encoded structure. This is consistent with previous prospective results using familiar events or segments, which have shown number-of-segments and predictability effects (Boltz, 2005; Waldum & Sahakyan, 2013). Importantly, the present results indicate that in addition to top-down knowledge, explicitly or incidentally learned event structure modulates duration judgments.

Task Demands in Prospective Paradigms

In the psychology of time literature, multiple judgment types and experimental designs have been used (Grondin, 2010). Some experimental designs may be more amenable to the use of memory than others. For example, the point at which judgments are requested after stimulus presentation—immediately or after a delay—may bias participants to rely on memory recollection (Zakay & Fallach, 1984). In particular, immediate judgments after each stimulus presentation surely involve a working memory representation of the stimuli (Grondin, 2005; Ogden et al., 2008), whereas delayed judgments—for example, after a few minutes of stimulus presentation or after the whole stimulus set has been processed—are likely to engage episodic memory, because the processed stimuli have to be committed to a longer-term memory store and cannot be forgotten before the next stimulus is processed. These observations suggest that the present prospective task design, as that in Boltz (2005), promoted the use of memory, and in particular, episodic memory representations that must be held until later in the experiment (Tulving & Thomson, 1973).

Other experimental protocols have also demonstrated an influence of episodic memory. Waldum and Sahakyan (2013) for example, requested prospective duration judgments after stimulus processing. The experiments presented different songs lasting for longer than 10 min in total while participants performed a secondary task. Such stimulus length cannot be held in a limited-capacity working memory (Cowan, 2001), and thus, the stimuli need to be committed to episodic memory in chunks or segments, as described in memory updating models (Radvansky & Copeland, 2006; Swallow, Zacks, & Abrams, 2009). As a result, participants provided judgments that correlated with the number of songs recalled. These results suggest a role of episodic memory in prospective judgments when the familiar nature and the long duration of the stimuli promote the use of encoded segmental structure to provide duration judgments.

Some experimental protocols surely need not rely on episodic memory, but most of them require stimulus encoding in some form, for example, in working memory, and therefore stimulus segmentation/chunking is an important factor to consider. Working memory experiments using seconds-long stimuli and duration judgments after each stimulus presentation have already highlighted the role of segmental structure. For example, many studies have shown that the number of sensory stimuli (segments) presented correlates positively with the judged duration (Buffardi, 1971; Poynter & Homa, 1983). Effects of stimulus predictability, which sometimes lead to opposite correlations with the number of segments, may depend on encoding difficulty and/or the segmentation strategies adopted. It is possible for example that stimuli that are easier to structure are judged shorter than stimuli that are harder to segment because of violations of prior expectations (Block et al., 2010; Liverence & Scholl, 2012; Meyerhoff et al., 2015; Schiffman & Bobko, 1974; Zakay et al., 1994). Differences in the role of event structure in prospective duration judgments may therefore stem from the structural analyses performed during encoding and its relation to prior knowledge, indicating a role for semantic memory and prior knowledge rarely considered in the literature.

These suggestions remain speculative and are unlikely to account for every single past or future finding. Surely, there could be tasks in which participants ignore stimulus content altogether and engage in other strategies like silently counting. Ultimately, the way in which participants approach the stimuli will depend on the type of information that they deem relevant given the instructions that they received before hand. Nevertheless, it appears that more nuanced distinctions need to be considered when discussing prospective paradigms: namely, whether they involve limited-capacity working memory for immediate judgment or episodic representations for later judgments, and in both cases, whether representations are formed through top-down expectations or bottom-up analysis, intentionally or incidentally.

Conclusion

The present findings suggest that newly formed episodic event memories play a role in prospective judgments: if participants have formed a memory of the events to be judged, they employ this representation to inform their judgments, even when they have also attended to duration. Specifically, fine-grained characteristics of the encoded event structure such as the number of segments and their similarity play a role in a prospective paradigm, just as they do in a retrospective paradigm. Moreover, even when participants have only attended to duration and did not explicitly encode content, they use some form of segmentation during stimulus processing that impacts on duration judgments. Previous prospective studies using a variety of protocols also indicate a role for encoded segmental structure, although this structure is likely influenced by prior expectations. Therefore, it appears that event structure, whether explicitly or incidentally encoded, is accessed and used in many prospective protocols.

References

Bartlett, F. C. (1995). *Remembering* (2nd ed.). Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511759185

Bellezza, F. S., & Young, D. R. (1989). Chunking of repeated events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 990–997. http://dx.doi.org/10.1037/0278-7393.15.5.990

Block, R. A. (1982). Temporal judgments and contextual change. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 530–544. http://dx.doi.org/10.1037/0278-7393.8.6.530

Block, R. A., & Gruber, R. P. (2014). Time perception, attention, and memory: A selective review. *Acta Psychologica, 149,* 129–133. http://dx.doi.org/10.1016/j.actpsy.2013.11.003

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica, 134,* 330–343. http://dx.doi.org/10.1016/j.actpsy.2010.03.006

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

Boltz, M. (1992). Temporal accent structure and the remembering of filmed narratives. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 90–105. http://dx.doi.org/10.1037/0096-1523.18.1.90

Boltz, M. G. (1995). Effects of event structure on retrospective duration judgments. *Perception & Psychophysics, 57,* 1080–1096. http://dx.doi.org/10.3758/BF03205466

Boltz, M. G. (2005). Duration judgments of naturalistic events in the auditory and visual modalities. *Perception & Psychophysics, 67,* 1362–1375. http://dx.doi.org/10.3758/BF03193641

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,* 487–502. http://dx.doi.org/10.1037/a0016797

Brown, S. W. (2008). Time and attention: Review of the literature. In S. Grondin (Ed.), *Psychology of time* (pp. 111–138). Bingley, UK: Emerald.

Brown, S. W., & Boltz, M. G. (2002). Attentional processes in time perception: Effects of mental workload and event structure. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 600–615. http://dx.doi.org/10.1037/0096-1523.28.3.600

Buffardi, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Perception & Psychophysics, 10,* 292–294. http://dx.doi.org/10.3758/BF03212828

Burt, C. D., & Kemp, S. (1991). Retrospective duration estimation of public events. *Memory & Cognition, 19,* 252–262. http://dx.doi.org/10.3758/BF03211149

Coll-Florit, M., & Gennari, S. P. (2011). Time in language: Event duration in language comprehension. *Cognitive Psychology, 62,* 41–79. http://dx.doi.org/10.1016/j.cogpsych.2010.09.002

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. http://dx.doi.org/10.1017/S0140525X01003922

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior, 11,* 671–684. http://dx.doi.org/10.1016/S0022-5371(72)80001-X

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

Faber, M., & Gennari, S. P. (2015b). Representing time in language and memory: The role of similarity structure. *Acta Psychologica, 156,* 156–161. http://dx.doi.org/10.1016/j.actpsy.2014.10.001

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin, 127,* 22–44. http://dx.doi.org/10.1037/0033-2909.127.1.22

Grondin, S. (2005). Overloading temporal memory. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 869–879. http://dx.doi.org/10.1037/0096-1523.31.5.869

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72,* 561–582. http://dx.doi.org/10.3758/APP.72.3.561

Hanson, C., & Hirst, W. (1989). On the representation of events: A study of orientation, recall, and recognition. *Journal of Experimental Psychology: General, 118,* 136–147. http://dx.doi.org/10.1037/0096-3445.118.2.136

Hays, W. L. (1981). *Statistics* (3rd ed.). New York, NY: Holt, Rinehart and Wiston.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491. http://dx.doi.org/10.1037/0033-295X.96.3.459

Liverence, B. M., & Scholl, B. J. (2012). Discrete events as units of perceived time. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 549–554. http://dx.doi.org/10.1037/a0027228

Meyerhoff, H. S., Vanes, L. D., & Huff, M. (2015). Spatiotemporal predictability alters perceived duration of visual events: Memento effect revisited. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 613–622. http://dx.doi.org/10.1037/xhp0000048

Ogden, R. S., Wearden, J. H., & Jones, L. A. (2008). The remembrance of times past: Interference in temporal reference memory. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1524–1544. http://dx.doi.org/10.1037/a0010347

Oppenheimer, D. M., & Kelso, E. (2015). Information processing as a paradigm for decision making. *Annual Review of Psychology, 66,* 277–294. http://dx.doi.org/10.1146/annurev-psych-010814-015148

Ornstein, R. E. (1969). *On the experience of time*. New York, NY: Penguin Books.

Poppenk, J., Köhler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1321–1330. http://dx.doi.org/10.1037/a0019900

Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition, 11,* 77–82. http://dx.doi.org/10.3758/BF03197664

Poynter, W. D. (1989). *Time and human cognition: A life-span perspective* (Advances in psychology, Vol. 59). Amsterdam, the Netherlands: Elsevier. http://doi.org/10.1016/S0166-4115(08)61045-6

Poynter, W. D., & Homa, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics, 33,* 548–560. http://dx.doi.org/10.3758/BF03202936

Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition, 34,* 1150–1156. http://dx.doi.org/10.3758/BF03193261

Schiffman, H. R., & Bobko, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology, 103,* 156–159. http://dx.doi.org/10.1037/h0036794

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134,* 207–222. http://dx.doi.org/10.1037/0033-2909.134.2.207

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General, 138,* 236–257. http://dx.doi.org/10.1037/a0015631

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373. http://dx.doi.org/10.1037/h0020071

Waldum, E. R., & Sahakyan, L. (2013). A role for memory in prospective timing informs timing in prospective memory. *Journal of Experimental Psychology: General, 142,* 809–826. http://dx.doi.org/10.1037/a0030113

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 61,* 569–587. http://dx.doi.org/10.1080/17470210701282576

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,* 1363–1374. http://dx.doi.org/10.1098/rstb.2001.0939

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28,* 979–1008. http://dx.doi.org/10.1207/s15516709cog2806_5

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138,* 307–327. http://dx.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,* 273–293. http://dx.doi.org/10.1037/0033-2909.133.2.273

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

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130,* 29–58. http://dx.doi.org/10.1037/0096-3445.130.1.29

Zakay, D., & Block, R. A. (1996). The role of attention in time estimation processes. *Advances in Psychology, 115,* 143–164. http://dx.doi.org/10.1016/S0166-4115(96)80057-4

Zakay, D., & Block, R. A. (1997). Temporal cognition. *Current Directions in Psychological Science, 6,* 12–16. http://dx.doi.org/10.1111/1467-8721.ep11512604

Zakay, D., & Fallach, E. (1984). Immediate and remote time estimation—A comparison. *Acta Psychologica, 57,* 69–81. http://dx.doi.org/10.1016/0001-6918(84)90054-4

Zakay, D., Tsal, Y., Moses, M., & Shahar, I. (1994). The role of segmentation in prospective and retrospective time estimation processes. *Memory & Cognition, 22,* 344–351. http://dx.doi.org/10.3758/BF03200861

Received April 11, 2016
Revision received October 30, 2016
Accepted November 10, 2016 ■