LUND
UNIVERSITY

DEPARTMENT OF PSYCHOLOGY

The Effects of Event Segmentation on Object Recognition Memory and Retrospective Temporal Judgements About Objects

Ebba Campbell

Saara Siintola

Bachelor’s Thesis, spring 2020

Advisor: Mats Dahl

Abstract

The world presents itself to our sense organs as a continuous stream of input, yet we appear to conceive experiences in terms of discrete events. According to event segmentation theory, this segmentation is an intrinsic and automatic component of human cognition. A considerable body of research provides evidence for this claim and suggests further that the way we segment events has consequences for memory: performance in recall and recognition tests is superior for information made available at event boundaries. The present study was set out to reproduce these previous findings and to further explore whether event boundaries are associated with an increase in the perceived duration of events. We hypothesized that event boundary objects would be recognised more effectively and judged retrospectively to have been shown for a longer period of time than comparable control objects. We conducted a repeated measures experiment with healthy adult participants which consisted of an encoding phase involving visual stimuli and a subsequent task on object recognition and temporal estimation. No statistically significant effects in support of either hypothesis were observed. Possible explanations for the lack of statistically significant results are the low statistical power of the study as well as shortcomings related to methodology.

Keywords: event perception, event segmentation theory, episodic memory, object recognition, temporal perception

Introduction

The way the world presents itself to us is fleeting in nature: the sensory inputs received by the receptors in our eyes, ears and other sense organs consist of continuous, dynamic activity. Despite this, we appear to conceive experiences in terms of discrete events which contain distinctive parts (such as a beginning and an end) and stand in hierarchical relations to other events. The activity of cooking a meal, for instance, could be described in terms of sub-events such as preparing the pre-course, preparing the main course and preparing the dessert, all of which can themselves be described as consisting of various sub-events (such as chopping vegetables, mixing ingredients in a bowl, seasoning, etc).

Our habit of describing sensory experiences in terms of discrete events does not in itself implicate that event segmentation would be a component of human perception. Rather than being a quality of perception and memory, the apparent chunking might just as well have to do with the way language is structured. However, a growing body of evidence from both behavioural studies and neuroimaging suggests that event segmentation is indeed a core component of perception which has consequences for both learning and memory (Zacks & Tversky, 2001; Zacks & Swallow, 2007). It appears that much like segmenting an object *spatially* into its parts is central for object recognition and understanding (Biederman, 1987), segmenting in *time* is central for understanding events.

In the literature concerning event perception, an event has been defined as “a segment of time at a given location that is perceived by an observer to have a beginning and an end” (Zacks & Tversky, 2001, p. 3). This definition is illustrating, but not exhaustive. The notion of an event can perhaps best be defined in terms of the wittgensteinian concept of family resemblance (Wittgenstein, 2009). Zacks, Speer, Swallow, Braver, & Reynolds (2007) point out that there are some features that events typically share (such as involving animate agents, being goal-oriented and ranging in time from a few seconds to several hours), but this does not mean that even highly atypical instances (such as the decay of a radioactive substance) could not be perceived and described as events. There exists, thus, no set of individually necessary and jointly sufficient conditions for something to count as an event. This definitional vagueness should pose no risk for the psychology of event perception, however: the ambiguity regarding event boundaries and taxonomy is actually analogous to the psychology of objects (Zacks et al., 2007).

Event Segmentation Theory

Based on findings in behavioural and neurological studies about event perception, Zacks & co proposed in 2007 a model called event segmentation theory (EST). According to the EST, the segmentation of ongoing activity into distinctive events is an intrinsic part of human cognition. This segmentation emerges from a perceptual processing stream whose input is a set of sensory representations (i.e. sensory data relayed to the cortex from our sense organs) and whose output is a set of perceptual predictions which allow us to anticipate the future state of our perceptual representations. In other words, raw sensory inputs are transformed in the perceptual stream to produce *"multimodal representations with rich semantic content, encoding information such as object identity and location, motion trajectories, and the identities and attitudes of other people*" (Zacks et al., 2007).

A central premise of event segmentation theory is that event perception is driven by so-called event models. Event models are multimodal working memory representations of "what is happening right now" which direct the processing in the core processing stream described above. They are informed by both changes in the present perceptual stream (bottom-up), and by event schemata which contain semantic information about the central features of previously encountered events (top-down) (Zacks, Tversky, & Gowri, 2001). Event models are relatively robust to transient variability in the incoming sensory input, which enables the perceptual constancy required to perceive an ongoing activity as an entity even when disruptions occur in the sensory input (Zacks et al., 2007).

Event models are also necessary for perceptual predictions to emerge from the perceptual processing stream. Perceptual predictions, on the other hand, enable us to not merely react to incoming stimuli, but to anticipate future states of affair and to plan our reactions accordingly. In other words, the ability to perceive continuous sensory stimuli as unified events allows us to predict things such as the future locations of objects in space (Zacks et al., 2007). The adaptiveness of this sort of ability for, for instance, an animal trying to avoid predators or catch prey should be evident. In a similar way, the ability to make event-based interferences about a person’s goals allows us to predict what they are going to do or say next.

The predictions we make about the future are of course not perfect – sometimes people and objects behave in ways which surprise us. EST suggests that the quality of perceptual predictions is continuously evaluated by an error detection mechanism. This mechanism functions by comparing the predictions we have made to the actual activity taking place as indicated by the input in the perceptual stream: The higher the correspondence between the two, the more successful the prediction. For the most part, the current event model is a sufficiently good fit for the current event and, as a result, perceptual prediction is effortless and accurate. However, when conceptual (top-down) or perceptual (bottom-up) features of the ongoing activity change too much, the contents of the current event model become less useful for perceptual prediction. This eventually leads to a transient increase in prediction errors and an updating of the event model (Zacks et al., 2007).

It is proposed that the updating of event models in the perceptual stream is a two-stage process. First, a gating mechanism reacts to the increase in prediction errors by (A) resetting the current event model and by (B) briefly increasing the influx of sensory inputs to the event models. These updates then lead into a decrease in prediction errors, which reduces the amount of input from sensory pathway back to normal and brings the event model back into a steady state (Zacks et al., 2007). In this way, the system alternates between short periods of change, subjectively experienced as event boundaries, and longer stable periods, subjectively experienced as ongoing events (Zacks et al., 2007; Sargent et al., 2013).

A tangible illustration of how the updating of an event model might occur is as follows: When watching a shop assistant at a clothing store fold shirts on a display table, first picking up one shirt, folding it and placing it on the table, then picking up another one, folding it and placing it on the table too, we might for a number of reasons expect that this pattern of behaviour will continue. First of all, continuing to pick up shirts and fold them would be maintaining a consistent movement pattern. Second of all, due to our previous experiences of seeing this activity take place, we know that that is how it usually goes. And lastly, we might make and inference about the shop assistant's goal (to fold all of the shirts). For as long as the shirt-folding activity continues, this prediction is successful. After the last shirt has been folded, however, anticipating the next move is likely to become more difficult, leading to a mismatch between prediction and perceptual input. This transient increase in prediction errors triggers the gating mechanism: the old event model is updated through resetting of the current event model and a quick increase in the power of sensory input. Influenced by both perceptual information and semantic information (event schemata), a new event model is eventually stabilised (an event model about walking towards the counter, perhaps). On a subjective level, this is experienced as an old event coming to an end and an as the beginning of a new event.

The error detection mechanism proposed by the EST has been questioned by one scientist in particular. Tauzin (2015) claims that such mechanism could not be in effect when a person is experiencing something for the first time. Tauzin himself proposes and has provided evidence for the possibility that even visual cues can trigger segmentation, particularly in unfamiliar situations. In a similar line of argument, (Khemlani, Harrison, & Trafton, 2015) point out that EST cannot explain how the error detection mechanism operates when different cues are competing against each other, or how an event could be represented neurocognitively. Admittedly, which cues are prioritized over others is not described in the theory. Therefore, as indicated Tauzin’s results, both predictability errors and changes in sensory cues are likely contributors to event segmentation.

Research Concerning Event Segmentation

Neurophysiological Evidence of Event Segmentation

Perceived event boundaries have since long been associated with specific patterns of neural activity, providing compelling support for the idea that event segmentation is, indeed, an intrinsic component of perception, rather than a property of the language used to describe events. In one study, an increase in pupil diameter – indicative of elevated perceptual processing – was correlated in time with perceived event boundaries (Swallow & Zacks, 2004). Research on the neural correlates of prediction errors have identified various brain regions particularly in the frontal cortex as candidates for the home of prediction errors (Dickinson & Schultz, 2000). In the framework of event segmentation theory, these structures should be responsible for signalling to the rest of the brain when a current event model should be updated to a new one. In one fMRI study, the perception of event boundaries in films was associated with increased activity in the posterior and frontal cortex which began before the identified event boundary and peaked after it, even for passive viewers (Zacks et al., 2001). Similar results were also demonstrated in an EEG study whose results indicated that the perception of event boundaries was correlated increased activity in the frontal and parietal lobes (Sharp, 2007).

Identification of Event Boundaries

Despite the neurophysiological evidence for event segmentation, what defines a break between one event and the next is ambiguous. Even for very simple events, such as placing an item a box, the boundaries could be defined in numerous ways — for example by reference to the picker's perceived goals, the motion trajectories of their body, an interaction between the picker's hand and the item, or some combination of these factors. Nevertheless, research indicates that the identification of event boundaries seems to be a case of "I know it when I see it": numerous studies have demonstrated a high agreement among different observers about the location of event boundaries in event narratives, and an even higher agreement within individuals across time) (Newtson & Engquist, 1976, Speer, Swallow, & Zacks, 2003, Kurby & Zacks, 2008). Research by (Tauzin, 2015) has further indicated that people segment events consistently even when the stimuli consists of abstract, unfamiliar stimuli consisting of flashing figures in a film clip; event boundaries were identified during both kind-related changes (an object being replaced by a new one), and kind-irrelevant changes (an object changing location).

One conclusion implied by the fact that events are experienced as hierarchical structures where one event consists of multiple sub-events (which consist of sub-events, and so forth), is that the processes involved in event segmentation occurs concurrently on a range of different time scales. In the event of baking a cake, for example, the whole baking process would occur on a relatively coarse time scale, whereas the sub-events of adding each different ingredient in the baking bowl would be seen as occurring on finer time scales. These phenomena are referred to in the literature as coarse-grained segmentation and fine-grained segmentation, respectively (Zacks et al., 2007).

If segmentation takes place on multiple time scales simultaneously, it should follow that variations in predictability also occur on both finer and coarser scales, and that the error signals triggering the resetting of an event model are adjusted to fit the appropriate grain (Zacks et al., 2007). As already described, increases in prediction errors are connected to the transitions from one event to the next. In the case of fine-grained segmentation, event models are hypothesized to react to and reset when smaller, relatively brief increases in prediction error occur. Correspondingly, coarse-grained representations are believed to require larger and more sustained increases in prediction error in order to update (Zacks et al., 2007).

Some evidence for the separation of fine- and coarse-grained segmentation on a processual level is provided by studies with patients who suffer from schizophrenia and frontal lobe lesions found that some such patients were selectively impaired in coarse-grained segmentation but experienced no difficulties in segmenting events on a finer time scale (Zalla, Verlut, Franck, Puzenat, & Sirigu, 2004; Sirigu et al., 1995). Other studies indicate that different cues may be emphasized by the error detection mechanism depending on the timescale at hand. In coarse-grain segmentation, an event boundary is typically detected when a conceptually or contextually driven change occurs, whereas fine grained segmentation appears to have a stronger connection to the perceptual (sensory) qualities of the ongoing activity (Zacks, Tversky, & Gowri, 2001).

Segmentation agreement — the degree to which an individual person's identifications of event boundaries corresponds to those of the group — has been found to effectively predict event memory, which indicates that normative segmentation is adaptive (Sargent, et al., 2013). Furthermore, the number of event boundaries a person perceives is predictive of better memory: several studies provide evidence that participants who indicate more event boundaries in a movie perform better in a subsequent recall test (Lassiter, Stone, & Rogers, 1988; Sargent et al., 2013). These results imply that event segmentation is a cognitive process whose effectiveness varies across individuals, with some of us being better at chunking activity into units that are efficient for the encoding and recollection of memories. This conclusion is further supported by studies indicating that segmentation agreement tends to decline with age. Young adults show more agreement over where an event boundary occurs than older adults do, with patients suffering from Alzheimer's being particularly limited in both event segmentation and later recollection of events (Zacks, Speer, Vettel, & Jacoby, 2006). Alzheimer's, along with other neuropathological conditions such as PTSD, Schizophrenia, and brain lesions, is supposed to make the perceptual organisation involved in event segmentation more difficult. Because of the difficulties in treating some of these conditions, new intervention approaches designed to enhance event segmentation ability might lead to improvement in life quality (Flores, Bailey, Eisenberg, & Zacks, 2017; Gold, Zacks, & Flores, 2017);. In a similar manner, the findings of event segmentation research might be potentially useful for developing new diagnostic tools for conditions associated with decreased event segmentation ability.

Event Segmentation and Memory Retrieval

According to event segmentation theory, our ability to recall a particular piece of information should be influenced by *when* during the event perception process the information was presented – more precisely, whether the information is originally perceived at an event boundary, or during an ongoing event. As the perceptual processing of information speeds up during event boundaries, even the memory encoding of said perceptions should be enhanced for information made available at these times. A myriad of research indicates that this is indeed so, and that both learning and memory can be improved by cueing event boundaries (Sargent, et al., 2013; Zacks, Kurby, Landazabal, Krueger, & Grafman, 2016).

Most memory research on event segmentation have used methods involving visual tasks. (Newtson & Engquist, 1976), for instance, found in one experiment that the detection of deleted sections in a film was more accurate for deletions made at event boundaries (breakpoints) than for deletions made within ongoing events (non-breakpoints). In another experiment in the same study, they found that the recognition accuracy during a recall test was consistently and highly significantly more accurate for items encountered during a breakpoint. a consistent and highly significant (Newtson & Engquist, 1976). In a similar manner, research by (Kurby & Zacks, 2008) indicates that memory for films depicting everyday activities (such as washing the dishes) is better for details which occurred at event boundaries. A series of experiments by (Swallow, Zacks, & Abrams, 2009) measured object recognition memory in particular, and found, in line with other research, that recognition accuracy was superior for objects which appeared during an ongoing event boundary compared to other objects. In addition to adult participants used in the vast majority of event segmentation research, visual paired comparison tests indicate that infants' memory is enhanced for objects made available at event boundaries, even before two years of age (Sonne, Kingo, & Krøjgaard, 2017).

One interesting application of these findings is implicated by research on the effects of commercial breaks in filmed narratives. A study by (Boltz, 1992) indicates that the highlighting of event boundaries with commercial breaks yields improved performance in subsequent recall and recognition tests for both information included in the film narrative and the temporal order of said information. Correspondingly, commercial breaks which obscured the natural event segmentation of the narrative by taking place in the middle of ongoing events led to significantly lower performance in recall and recognition tests. Similar results have been demonstrated also for tasks involving the learning of a procedure. Highlighting the hierarchical structure of the procedure facilitated learning, whereas obscuring or misrepresenting it impaired memory and learning (Zacks & Tversky, 2003). These findings could be of interest to, for instance, the film industry and educational services.

In addition to the visual modality, individual studies exist which have investigated the effects of event segmentation on memory using other kinds of stimuli, such as language (Speer, Zacks, & Reynolds, 2004) or music (Sridharan, Levitin, Chafe, Berger, & Menon, 2007). While these results indicate that event segmentation occurs across different modalities, the body of research is much less extensive.

Temporal Perception of Events

The subjectively experienced duration of events often differs from objective time measurements. We tend to, for instance, overestimate the length of short events and underestimate the length of long events (Lejeune & Wearden, 2009), and our estimations on the duration of events which have an irregular or unpredictable tempo will often be less accurate (Boltz M. G., 1998). Perhaps unsurprisingly, emotions involved in an event can distort our retrospective judgements about event duration (Fayolle, Droit-Volet, & Gil, 2014). One example of this is that length of stressful events, such as being involved in a car crash or witnessing a crime, is consistently overestimated (Block & Gruber, 2014).

One central finding in the psychology of time is that present-time, or *prospective* temporal judgments are differ both theoretically and empirically from past-time, or *retrospective* judgments (Block R. A., 1990). Experiments utilising retrospective duration judgements have found that in situations where participants are tasked to estimate the duration of an event numerically (in minutes or seconds), increases in contextual changes and cognitive load are associated with longer estimated durations (Block & Reed, 1978; Block, Hancock & Zakay, 2010).

In spite of the clear connection between the perceptual qualities of an event and retrospective judgements about its duration, the effects of event segmentation on retrospective temporal judgements are largely unexplored. Considering that event segmentation is triggered by contextual changes and/or increases in the influx of sensory input, both of which have been associated with longer retrospective duration judgements, it would seem intuitive that objects presented at event boundaries would be judged as having been shown for a longer time than comparable non-event boundary objects.

Though not explicitly focused on event segmentation, some research on the temporal perception of events provides evidence for this intuition: predictable stimuli is experienced as more compressed in time compared to unpredictable stimuli (Pariyadath & Eagleman, 2007). Another study by Ezzyat and Davachi has provided evidence that event segmentation can influence retrospective temporal judgements, although the focus in the study was on items' perceived location in time relative to each other, rather than their absolute duration. The participants in the study first observed a set of images of items paired with different contexts, and then indicated on a Likert scale (1 = *very close* to 5 = *very far*) how close to each other in time certain objects had been. The results indicated that items separated with a contextual shift – an event boundary - were perceived as being farther apart in time (Ezzyat & Davachi, 2014).

Aim of the Present Study

The purpose of this study is to explore the influence of event segmentation on object recognition memory (whether we can duplicate the findings established by previous studies which indicate that memory is enhanced for even boundary objects) and the temporal perception of said objects.

While the influence of event segmentation on recall memory (*that* you saw an object) has been widely researched, few studies have been done to explore the influence of event segmentation on temporal perception (*for how long* you saw an object). Considering the evidence that increased perceptual processing at event boundaries enhances object recall and that the perceptual qualities of an event can influence temporal estimations concerning said event, it would not seem strange if event segmentation had a consistent effect on our temporal memory as well. As explained in the previous chapter, some research has already been done which suggests that the perceived temporal distance between different objects is larger for object pairs separated by an event boundary (Ezzyat & Davachi, 2014).

We are interested in finding out whether the location of a stimulus in relation to an event boundary influences the perception of how long the stimulus itself was perceived. To our knowledge, no previous study has examined this exact question. ¨

In order to assess the effects of event segmentation on object recognition and the retrospective temporal judgements of said objects, we tested the ability of people to recognise object pictures shown to them in a prior testing phase of the experiment, and their ability to estimate the temporal duration (in seconds) said objects had been visible. We hypothesised that people would be better at recognising object pictures which had been shown during event boundaries (event boundary objects) than object pictures shown during ongoing events (control objects). To induce as high level of segmentation agreement as possible between the participants, we operationalised event boundaries as a sudden change in an otherwise unchanging contextual location (background picture) of an object.

Building upon the aforementioned research indicating that memory performance is enhanced for unpredictable and novel stimuli and that event segmentation affects the perceived relative location of objects in time, we hypothesised further that event boundary objects would retroactively be judged as having longer duration than control objects.

Method

Participants and Ethics

34 neurologically healthy participants (18 females and 16 males) between the ages of 21 and 81 (M = 30.58, SD = 24.75) took part in the experiment. The participants were recruited through convenience sampling (classmates, work colleagues, friends, and family members of ours) from around the Øresund Region in Sweden and Denmark. Each

participant had either normal or corrected vision, and normal colour vision. One participant had to be excluded from the analysis due to a technical malfunction during the experiment, and an additional seven participants due to coding errors, resulting in 26 participants (12 women and 14 men) being included in the analyses. No money was offered as compensation for participation, but chocolate was available for the participants to enjoy after their participation in the study.

The recruitment and experimental procedures used in the study were approved in advance by our thesis advisor and course coordinator and complied with the ethical guidelines of Lund University and Swedish law. Each time before beginning the experiment, the purposes and method of the experiment, and the right to discontinue at any time were explained to the participant who then read and signed a consent form (see Appendix A) containing the same information. The participants were also encouraged to voice any concerns or questions they might have about the experiment. No questions or concerns were raised, however, and everyone decided to proceed with the experiment. After signing the consent form, participants were requested to fill a background information sheet (see Appendix B) asking about their vision, age, gender, and education level to make sure all participants had normal vision and to enable the statistical testing of interaction effects for age, gender, or education level. The background data was pseudonymised (participant ID numbers were used to experiment data with the background sheets) and processed, as all data used in the experiment, confidentially and in accordance with Lund University’s guidelines and Swedish data protection legislation.

Design, Materials, and Procedure

The experimental design used in the study is an example of a complete repeated measures design, as it consisted of two conditions (experimental and control) both of which were administered to each subject several (four) times. Possible practise effects were balanced both within and across the participants by randomising the order of these blocks. The alpha level was set a priori at .05 for all statistical tests, as this is conventional and yields statistical power at a level where differences can be detected even for smaller sample sizes.

The testing was done in a single, 10-20 minutes long testing session (per participant) with two main phases (an encoding phase and a testing phase), as well as a short distraction task administered in between. In order to minimise external distractions, each session took place in a silent room specifically reserved for the experiment. We made sure that both instructions and the actual tasks were presented in an identical manner for each participant by using the program E-Prime 3.0 (standard 16:9 display ratio with a 1366 x 768 resolution on a Fujitsu Lifebook s751 laptop) to administer them. Other materials used in the experiment included a pen and a paper which the participants used during a distraction task taken between the encoding and testing phases.

The stimuli we used consisted of scene photographs and object images taken from the Internet-based image bank, Bradylab (https://bradylab.ucsd.edu). To ensure sufficient homogeneity of the stimuli, all pictures we chose were emotionally neutral photographs with a good resolution, depicting real-life environments and objects (see Figure 1 for an example of a pair of object pictures).

Figure 1

A pair of objects

Note. Example of a pair of objects used as stimuli in the experiment.

In the encoding phase, the participants watched a short visual presentation consisting of eight successive blocks, the order of which was counterbalanced within each participant. In each block, seven pictures of different but semantically related objects were shown successively on top of a thematically compatible background image (e.g. seven nature-related objects on top of a nature picture, or seven kitchen-related objects on top of a kitchen picture). The presentation of individual objects was designed to trigger event segmentation on a finer scale (each object representing a fine-grain event), whereas the presentation of background photos was designed to trigger event segmentation on a coarser grain (each background photo representing a coarse grain event). The purpose of the thematical compatibility between object pictures and background pictures was to increase the ecological validity of the events (the objects were a natural match for the background) and to make it even easier to perceive the presentation of background photos as meaningful events despite the fact that the background photos themselves were just static pictures.

For the purposes of the study, we were only interested in measuring the participants' performance concerning the middle object of each block. These middle objects (objects number 4) were always shown for a duration of 4 seconds, whereas the objects preceding and succeeding it (objects 1-3 and 5-7) were shown for 2 seconds each. The purpose of this temporal variation was to prevent ceiling effects in the upcoming temporal recall test, while keeping the duration of each object position (e.g. the critical middle objects) constant across blocks to inhibit confounding influences. The main reason for using events of such short duration was that we wanted to be able to test each participant in both conditions (experimental and control) several times, while also keeping the total duration of the experiment relatively short.

The experimental and control conditions were designed as follows: for a half of the blocks (blocks of type A) the background image remained unchanged for the duration of the block: each of the seven object pictures were shown on top of the same, unchanging background picture. These blocks constituted the control condition of the experiment. For the other half of the blocks (blocks of type B) the background image was changed to a new (but still thematically compatible) background image in the middle of the block, so that the change always co-occurred with the presentation of the middle-object (see Figure 2 for an example).[1] These blocks constituted the study's experimental condition.

As the middle objects in blocks of type B were always associated with a contextual shift in the background (constituting an event change), whereas no such change occurred in blocks of type A (constituting and ongoing event), the design ensured that the only relevant difference between the critical middle objects was whether or not their presentation was associated with a (coarse grain) event boundary.

Figure 2

Background change

Note. Example of the type of shift in background image (event boundary) that was used in blocks of type B (the experimental condition). The butterfly represents the critical middle object in this block.

---

[1] No flickering occurred in the object picture during the background change; the way the middle object itself was presented was thus identical for both block types, A and B.

To prohibit order and practice effects, the order of the blocks was randomised using the built-in randomisation function of E-Prime so that any block could appear in any position to any participant. Furthermore, the order of the object images shown within the blocks was also randomised. However, due to technical limitations of E-Prime, this had to be done manually for blocks of type B after every participant had completed the experiment: we re-defined the objects' relative positions in the program after every participant, so that each of the seven objects in a block would appear in the middle position (or any other position) equally often. In order to ensure successful completion of this task, we used a pre-written aide-memoire specifying the objects' positions for each upcoming participant

After the encoding phase, the participants completed a short distraction task (15 seconds to write down movie titles that begin with the letter "L"). The only purpose of this task was to clear the participants' working memory of any contents related to the encoding phase, preventing rehearsal effects in the following recognition test.

In phase two, the recognition test, the participants were presented with subsequent object images (N=112) and asked to indicate by pressing a key whether the images were familiar from the encoding phase or not. The order of these images was randomised with E-Prime's own randomisation function to counterbalance possible recency and primacy effects across participants. A half of the images had been shown in the encoding phase, whereas the other half were novel images resembling the ones seen during encoding (see Figure 1 for an example of the types of objects used). The participants were also asked to rate their confidence in each familiarity evaluation on a Likert scale (1 = *not at all confident*, 2 = *somewhat confident*, 3. = *confident*, 4 = *very confident*, and 5 = *extremely confident*). By complementing the recognition task with a confidence rating, we made it possible to analyse data on confidently recognised objects separately from that of all recognised objects. The difference is of importance, because some objects were likely indicated as familiar as a result of guessing or relatively unsure judgement.

If a participant had indicated that a particular image was familiar, they were further requested to estimate the duration (in full seconds) that image had been visible on screen, on a scale from 1 to 8. This way of measuring perceived duration is, admittedly, rather rigid and insensitive, but programming a more dynamic measure within E-Prime proved too difficult within the given time frame. In order to avoid cueing the participants to adjust their estimations downwards, we set the longest possible estimation to be twice as long as the actual duration (which was always 4 seconds).

Results

The mean recognition percentage of the critical middle objects was 75% for event boundary objects (blocks of type B) and 71% for control objects (blocks of type A). A two-way mixed between-within subjects ANOVA was conducted to assess the impact of event segmentation on object recognition in general as well as possible interaction effects for gender No statistically significant interaction effect between gender and experimental condition (event boundary objects or control objects) was found in the analysis, Wilks' Lambda = .95, F (1, 24) = 1.18, *p* = .29, partial eta squared = .05. There was also no significant main effect for type of object, Wilk's Lambda =.99, F (1, 24) = 1.7, *p* = .68, partial eta squared =.007. It was self-evident that the results would be similar even when only confidently recognised objects, as indicated by the participant's having reported that they were either "very confident" or "extremely confident" in their answer, were included in the analysis: the average percentage of confident recognitions was, namely, identical for both event boundary objects and control objects at 52%.

Another two-way mixed between-within subjects ANOVA was conducted to assess the impact of event segmentation on retrospective temporal judgements about object images used in the experiment. Even here, the interaction between gender and experimental condition was not significant, Wilk's Lambda = .98, F (1,21) = .41, *p* = .53, partial eta squared = .02. The main effect for experimental condition was also not statistically significant, Wilk's Lambda = .98, F (1,21) = .48, *p* = .50, partial eta squared = .22. The main effect for gender was significant at the p < .05 level, F (1,21) = 4.59, p = .04, partial eta squared = .18, suggesting that women estimated the time an object had been presented during the encoding phase as longer than men, regardless of object type. In this case, this also means that women outperformed men: the mean temporal estimation for event boundary objects was 3.26 seconds (SD = .89) for women and 2.50 seconds (SD = 1.45) for men, whereas the mean temporal estimation for control objects was 3.64 seconds (SD = 1.34) for women, and 2.51 seconds (SD = 1.12) for men. The total mean estimation was 2.9 seconds (SD = 1.27) for event boundary objects and 3.05 seconds (SD = 1.36) for control objects.

Discussion

Our purpose with the present experiment was to investigate whether people's performance in an object recognition test would, as a considerable body of previous research suggests, be superior for objects whose presentation co-occurred with an event boundary compared to control objects presented during ongoing events. Furthermore, we wanted to find out whether event boundary objects would retroactively be judged as having been shown for a longer absolute time in seconds than control object; both the theoretical underpinnings of event segmentation theory, general research on the temporal perception of events, and the aforementioned study by (Ezzyat & Davachi, 2014) on the effects of event segmentation on the perceived relative temporal location of objects suggest that this might be the case.

Background information about age, education level, and gender were collected from the participants which allowed testing for possible interaction effects. These were for interest, since research indicates that segmentation ability declines with age (Zacks et al., 2006) and that episodic memory in general is worse for older adults (Nyberg, Bäckman, Erngrund, Olofsson, & Nilsson, 1996), people with higher education level outperform people of lower education levels (Angel, Fay, Bouazzaoui, Boudouin, & Isingrini, 2010) and women outperform men in tasks measuring episodic memory (Herlitz, Nilsson, & Bäckman, 1997). Due to the limitations of the sampling used in the study, over a half of the participants were under 30 years of age and had specified upper secondary school as their highest completed education level. Therefore, the only background variable we included in the analysis was gender.

Neither main effects for object type nor interaction effects between object type and gender were found for either object recognition or temporal estimation. The object recognition data indicated that event boundary objects were recognised more times than control objects, which is in line with our first hypothesis, whereas the data for temporal estimation indicated that control objects were perceived as having longer duration, which contradicts our second hypothesis. However, since neither result was even close to approaching significance, no conclusions can be drawn on the basis of these tendencies in either direction. The failure to observe the predicted effects could, in principle, be taken as evidence inconsistent with the hypotheses, but due to the limitations of methodology and statistical power of the experiment, it is questionable whether such interpretation could be justified.

One somewhat interesting result of the study is the finding of a significant main effect of gender on temporal estimation (women outperformed men). Though we were not concerned with measuring the effects of gender, the fact that this result is in line with previous studies on episodic memory, constitutes some (limited) evidence in support of the presumption that our test did, indeed, measure performance on episodic memory.

We chose a repeated measures design because having each individual participate in both experimental and control conditions reduces inter-individual variability, allowing for a higher internal validity of the experiment. Practice effects during the encoding phase where participants learn to predict what happens next (potentially influencing the characteristics of their event segmentation activity) were deemed likely, but relatively easy to control for by administering each condition several times (a so-called complete design) and by randomising the order of the blocks.. This design is known for being well-suited for detecting small differences in cognition (Shaugnessy, Zechmeister, & Zechmester, 2009, p. 227).

Another potential threat to the internal validity of the experiment is the loss of respondents due to technical malfunctions and coding errors. The limitation of the experiment program which compelled us to manually perform the randomisation of objects within blocks of type B resulted in a program shutdown once, forcing us to exclude one participant from the study. An additional seven participants in total had to be excluded from the study because their data indicated that the program had not functioned as intended. These attritions are likely not a huge problem for the internal validity of the study, however, since there is no particular reason to think that the occurrence of these issues would have been selective.

A further technical error in the coding of a block of type B led to the exclusion of that block for all participants. In order to correct for the created asymmetry created by this procedure, one randomly selected control block was removed as well. This transformed the study’s design into a 3x3 design, where only six of the eight original blocks were used for every participant. Due to the efficient randomisation of the order of the blocks, possible confounding influences created by the faulty block during the encoding phase should not affect the results of the study in a biased way. Nevertheless, this removal of 25% of the blocks used in the experiment meant that we had to conduct the analyses with only 75% of the original data. This could be a substantial factor in that no statistically significant results were obtained for the effects of event segmentation on either object recognition or temporal estimation.

The external validity of an experiment can be improved by using random sampling to select participants and by conducting the experiment in a natural setting, neither of which were true for our case. It is possible that interaction effects between biases in our sampling and the independent variable (event boundary vs. control) influenced the results of the study. On the other hand, the use of convenience sampling is generally less problematic for studies like ours which aim to evaluate intrinsic and universal components of human cognition, as opposed to studies evaluating, for instance, attitudes or personality which tend to vary a lot between individuals and groups. With that being said, the fact that only 26 of our original 34 participants could be included in the analyses means that the influence of extraneous confounding factors specific to particular participants was necessarily heightened. This is not optimal, although the number is likely sufficient for 80% power in studies using a repeated measures design (such as ours) to detecting large or perhaps even medium (d = .50) effect sizes (Aron, Aron & Coups, 2014).

Both the setting of the experimental and the operationalisation of the variables were highly controlled and laboratory-like. The experiment took place in a quiet room with both the procedure and presentation of the stimuli being standardised, and the probability of segmentation disagreement among participants was minimised through the use of stimuli where periods of sustained, unchanging presentation of still images alternated with total and complete changes of the stimuli being shown. The strength of this method is that it minimised the risk that segmentation will happen at different points in time for different people, an important factor to consider in a situation like ours where the lack of resources made it impractical to conduct another experiment prior to the primary one to find out where in the stimuli participants identified event boundaries. As explained previously, event segmentation is triggered by increases in prediction error or changes in sensory stimuli, but for still images, neither should occur. Thus, the only natural segmentation points in the stimuli we used were the intended points of segmentation: changes in the background images (coarse grain) and object images (fine grain). The flipside of this is the compromising of ecological validity: the typical event we encounter (be it in real life or e.g. within a movie) involves dynamic changes. Even though there is evidence that event segmentation occurs even for non-natural stimuli involving flashing objects on a screen (Tauzin, 2015), the segmentation of unnatural events might differ qualitatively from that of natural events. Caution is therefore warranted when drawing conclusions about event segmentation in general on the basis of data involving highly unnatural events.

Another, perhaps more serious risk for the construct validity of the experiment was related to the way temporal perception was measured. In the chosen method, participants would indicate their experienced duration of an object by choosing one of eight alternatives, ranging from 1 to 8 seconds. This way of measuring temporal perception is rather coarse, forcing the participants to choose round their experienced time to the nearest estimate. The offering of pre-selected alternatives is also likely to bias the participants estimates, e.g. through effects such as anchoring (Tversky & Kahneman, 1974) and the favouring or avoidance of extreme scores (Messick, 1962). These problems could have been solved e.g. by using a tactile approach where the participant hold down a key for as long as they think they saw a particular image. The reason such method was not used despite of its many advantages is that it proved too difficult to program in E-Prime. Be that as it may, the unsatisfactory operationalisation of temporal perception may be a central factor in our failure to detect even trends consistent with our temporal estimation -hypothesis.

One major improvement which could have helped to avoid many threats to the present experiment’s validity is the conducting of a pilot study to evaluate the suitability of the chosen stimuli and methods. Most importantly for our case, the unnecessary exclusion of participants due to coding errors could easily have been avoided if the problems had been discovered prior to the actual experiment.

It would also be useful to include a practice task in the experimental procedure to make sure each participant had understood the task as intended - e.g. the approximate level of how similar novel objects could be to the familiar ones. We used pairs of similar looking objects as stimuli in the experiment, but since we asked our participants to evaluate individual objects ("do you remember seeing this object?"), not compare object pairs ("which one of these objects did you see?"), this was actually unnecessary. We were, after all, interested in measuring differences in the recognition event boundary objects vs. control objects, not familiar vs. similar looking new objects. There was no harm in having some of the "filler objects" in the recognition task be similar to the objects we were interested in, but also no particular advantages, other than perhaps demonstrating to the participants that for successful recognition, it was important to recognise the objects on some level of detail (as novel objects could be of the same type as familiar ones).

Future Research

Since no statistically significant results or even near-significant trends were observed in support of either of our hypotheses, one logical possibility for future research would be to simply correct the shortcomings of the present study and see whether the results change. In our view, the research questions and hypotheses concerning temporal estimations in particular would be worthy of examining, regardless of our results. There is no reason to expect that a study otherwise similar to the present one, but with more participants and a more sensitive way of measuring temporal estimations could not produce significant results.

Inspired by the realisation that the use of object pairs was unnecessary for the purposes of our study, one future alteration to our design could be to have the participants compare two objects (one familiar and one novel) to each other in the recognition phase. This would make it possible to measure whether the level of detail in which event boundary objects are encoded and recalled differs from that of control objects. For such an experiment to successful in avoiding ceiling effects, the objects pairs used as stimuli would likely have to resemble each other more closely than the object pairs we used, as the success rate for our (presumably more difficult) single object recognition task was over 70% for both event boundary objects and control objects.

References

Angel, L., Fay, S., Bouazzaoui, B., Boudouin, A., & Isingrini, M. (2010). Protective role of educational level on episodic memory aging: an event-related potential study. *Brain and Cognition,* 312-323.

Aron, A., Aron, E., & Coups, E. (2014). *Statistics for Psychology.* Harlow: Pearson Education Limited.

Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review, 94*(2), 115-147.

Block, R. A. (1990). *Cognitive Models of Psychological Time.* Hillsdale: Lawrence Erlbaum Associates, Inc.

Block, R. A., & Reed, M. A. (1978). Remembered duration: Evidence for a contextual-change hyphothesis. *Journal of Experimental Psychology: Human learning and memory, 4*, 656-665.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgements: a meta-analytic review. *Acta Psychologica, 134*, 330-343.

Block, R., & Gruber, R. P. (2014). Time perception, attention and memory: a selective review. *Acta Psychologica, 149*, 129-133.

Boltz, M. (1992). Temporal Accent Structure and the Remembering of Filmed Narratives. *Journal of Experimental Psychology, Human Perception and Perofromance*, 90-105.

Boltz, M. G. (1998). Task predictability and remembered duration. *Perception & Psychophysics, 60,* 768-784.

Dickinson, A., & Schultz, W. (2000). Neural coding of prediction errors. *Annual Review of Neuroscience, 23*, 473-500.

Ezzyat, Y., & Davachi, L. (2014). Similarity Breeds Proximity: Pattern Similarity within and across Contexts Is Related to Later Mnemonic Judgments of Temporal Proximity. *Neuron*, 1179-1189.

Fayolle, S., Droit-Volet, S., & Gil, S. (2014). Emotion and Time Perception: Effects of Film-induced Mood. *Procedia - Social and Behavioral Sciences, 126*(21), 251-252. doi:https://doi.org/10.1016/j.sbspro.2014.02.399

Flores, S., Bailey, H. R., Eisenberg, M. L., & Zacks, J. M. (2017). Event segmentation improves event memory up to one month later. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(8), 1183-1202. doi:doi: 10.1037/xlm0000367.

Gold, D. A., Zacks, J. M., & Flores, S. (2017). Effects of cues to event segmentation on subsequent memory. *Cognitive Research: Principles and Implications, 2*(1). doi:10.1186/s41235-016-0043-2.

Herlitz, A., Nilsson, L.-G., & Bäckman, L. (1997). Gender differences in episodic memory. *Memory & Cognition*, 801-811.

Khemlani, S. S., Harrison, A. M., & Trafton, G. J. (2015). Episodes, events and models. *Frontiers in Human Neuroscience, 9*(590), 1-13.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 72-79.

Lassiter, D. G., Stone, J. I., & Rogers, S. L. (1988). Memorial consequences of variation in behavior perception. *Journal of Experimental Social Psychology*, 222-239.

Lejeune, H., & Wearden, J. H. (2009). Vierordt's The Experimental Study of the Time Sense (1868) and its legacy. *European Journal of Cognitive Psychology, 21*(6), 941-960. doi:https://doi.org/10.1080/09541440802453006

Messick, S. (1962). Response style and content measures from personality inventories. *Educational and Psychological Measurement*, 41-56.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, 436-450.

Nyberg, L., Bäckman, L., Erngrund, K., Olofsson, U., & Nilsson, L.-G. (1996). Age differences in episodic memory, semantic memory, and priming: relationships to demographic, intellectual, and biological factors. *Journal of Gerontology: Psychological Sciences*, 234-240.

Pariyadath, V., & Eagleman, D. (2007). The Effect of Predictability on Subjective Duration. *Plos One*. doi: https://doi.org/10.1371/journal.pone.0001264

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., . . . Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, 241-255.

Sharp, R. M. (2007). Electrophysiological correlates of event segmentation: how does the human mind process ongoing activity? *Proceedings of the Annual Meeting of the Cognitive Neuroscience Society* (p. 235). New York: Cognitive Neuroscience Society.

Shaugnessy, J. J., Zechmeister, E. B., & Zechmester, J. S. (2009). *Researh Methods in Psychology*. New York: McGraw-Hill Higher Education.

Sirigu, A., Zalla, T., Pillon, B., Grafman, J., Agid, Y., & Dubois, B. (1995). Selective imaprments in managerial knowledge following pre-frontal cortex damage. *Cortex, 31*, 301-316.

Sonne, T., Kingo, O., & Krøjgaard, P. (2017). Bound to remember: Infants show superior memory for objects presented at event boundaries. *Scandinavian Journal of Psychology*, 107-113.

Speer, N. K., Swallow, K. M., & Zacks, J. M. (2003). Activation of human motion processing areas during event perception. *Cognitive, Affective & Behavioral Neuroscience, 3,* 335-345.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2004). Perceiving Narrated Events. *Proceedings of the Annual Meeting of the Cognitive Science Society, 26.*

Sridharan, D., Levitin, D. J., Chafe, C. H., Berger, J., & Menon, V. (2007). Neural dynamics of event segmentation in music: converging evidence for dissociable ventral and dorsal networks. *Neuron, 55*(3), 521-532. doi:10.1016/j.neuron.2007.07.003.

Swallow, K. M., & Zacks, J. M. (2004). Hierarchical groupings of events revealed by eye movements. *Abstr. Psychonomic Society, 9,* 81.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, 236-257.

Tauzin, T. (2015). Simple visual cues of event boundaries. *Acta Psychologica*, 8-18.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 1124-1131. doi: 10.1126/science.185.4157.1124

Wittgenstein, L. (2009). *Philosophical Investigations.* (G. E. Anscombe, P. M. Hacker, & J. Schulte, Trans.) Chichester: Wiley-Blackwell.

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, 80-84.

Zacks, J. M., & Tversky, B. (2001). Event Structure in Perception and Cognition. *Psychological Bulletin*, 3-12.

Zacks, J. M., & Tversky, B. (2003). Structuring information interfaces for procedural learning. *Journal of Experimental Psychology, 9*(2), 88-100. doi:https://doi.org/10.1037/1076-898X.9.2.88

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., . . . Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651-655. doi:https://doi.org/10.1038/88486

Zacks, J. M., Kurby, C. A., Landazabal, C. S., Krueger, F., & Grafman, J. (2016). Effects of penetrating traumatic brain injury on event segmentation and memory. *Cortex, 74,* 233-246. doi:10.1016/j.cortex.2015.11.002

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.

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychological aging*, 466-482.

Zacks, J. M., Tversky, B., & Gowri, I. (2001). Perceiving, Remembering, and Communicating Structure in Events. *Journal of Experimental Psychology: General, 130*(1), 29-58.

Zalla, T., Verlut, I., Franck, N., Puzenat, D., & Sirigu, A. (2004). Perception of dynamic activity in patients with schitzophrenia. *Psychiatry Research, 128*(1), 39-51.

Appendix A

Consent form

Consent for participation in memory research

I volunteer to participate in a study conducted as a part of a bachelor's thesis project by two psychology students at Lund University. The purpose of the research project is to study how the human memory functions: its results are not indicative of intelligence or other such cognitive abilities.

My participation in the study is voluntary, and I may withdraw and discontinue my participation at any time without penalty.

The procedure involves observing a presentation on screen and answering to survey questions about the presentation afterwards. The study will take approximately 15-20 minutes and contains no explicit or upsetting material that would be likely to lead to lead to elevated levels of stress, anxiety, or similar psychological harm.

My responses will be confidential and no information about names, personal number, emails, or other such identifying factors will be collected in the survey; all data will be treated anonymously.

The results of this study will be used for scholarly purposes only.

I have read the above information and had my questions answered to my satisfaction. I am at least 18 years old.

_________________________ ______________________
My signature Date

____
My printed name

Appendix B

Background information sheet

Background information for a study about memory at Lund University

The information in this form will be used for the purposes of statistical analyses only and will be treated anonymously and confidentially.

Age: ________

Gender:
( ) woman
( ) man
( ) other

Highest level of education completed so far:
( ) no education
( ) basic/elementary education
( ) high school education
( ) bachelor's degree
( ) master's degree
( ) PhD or higher

Vision:
( ) I have normal or corrected (glasses/contact lenses) vision
( ) I am colour blind or have other such aberration of vision

Participant Nr ______________ Study completed at (time/date) _________________