Event Segmentation from Working Memory Load Manipulation

Sunjae Shim

Advisor: Dr. Christopher Baldassano

Columbia University

Abstract

The way that people divide up an experience into separate events can have long-term impacts on the way that it is remembered. According to a leading theory of event segmentation, an event boundary is perceived when an individual makes an error in their prediction about what will happen next. However, recent studies have raised doubts about whether an event boundary is created only when a prediction error occurs. We tested whether event boundaries can be triggered even without prediction error by manipulating working memory demands. In experiment 1, participants were shown a series of random images along with an image of a clock that indicated the beginning of a new segment and performed a task that required remembering images within the current segment. The clocks were manipulated across condition to either create a prediction error at the start of a new segment or to create a predictable transition. Then, participants completed a temporal order memory task to identify event boundaries in long-term memory, since previous research has shown that order judgements can be made with higher accuracy for pairs of items within the same event. We found that this boundary-related memory effect was present in both the predictable and unpredictable event boundary conditions. In experiment 2, participants either completed a working memory reset task based on predictable changes in background color or a working memory task without event structure as a control. A difference in temporal order memory accuracy for within and between segment items was found for the background condition but not for the control condition. This supports the hypothesis that active working memory updating can drive event boundary perception in the absence of a prediction error.

Event Segmentation from Working Memory Load Manipulation

Many of the experiences we go through in everyday life can be divided into a sequence of discrete events. For example, when asked to recount your experience of going to work in the morning, you could respond by listing the separate events of making your bed, taking a shower, making breakfast, and walking to the bus station. Segmentation of experiences seems to be natural for humans as people are able to identify event boundaries, or the beginning and end of each event, even during passive viewing of events or narrative film (Zacks et al., 2001; Ben-Yakov & Henson, 2018). People identify event boundaries at similar points based on a variety of features such as temporal shifts (Speer & Zacks, 2005), changes in location (Magliano, Miller, & Zwann, 2001), and changes in character's goals (Zacks, Speer, & Reynolds, 2009; Magliano et al., 2014). Event segmentation has been identified behaviorally through longer processing times at event boundaries; Radvansky and Copeland (2010) found that people read narrative stories slower at points of temporal shifts, possibly due to increased mental effort spent on updating the event model.

Event Segmentation Theory

How do people perceive event boundaries? Zacks and colleagues (2007) suggested an event segmentation theory which posits that event boundaries are perceived when a prediction error occurs. According to this theory, people hold mental representations, or event models, of their experiences in working memory. Based on perceptual input from the outside world and the current event model, predictions about what will happen next are created. When perceptual input and predictions are well matched, the event models are stable, but when an unexpected change occurs, their predictions no longer match the information gathered and this creates a prediction error. At this point, an event boundary is marked and the current event model is updated to incorporate the new information.

Several studies have investigated this model of understanding event perception. Zacks and colleagues (2011) asked participants to watch a naturalistic video clip and make predictions about what would happen, and found that moments where participants reported to have difficulty predicting accurately corresponded with subjective boundaries. Also, studies have identified that event boundaries often correspond with moments where there is a big change in perceptual features such as motion or situational features, which would correlate with points where prediction errors increase (Zacks, 2010; Hard, Tversky, & Lang, 2006; Zacks, 2004, Zacks, Speer, & Reynolds, 2009).

Event Perception and Memory

According to the event segmentation theory, working memory could be affected by event boundary perception. Speers and Zacks (2005) found that after an event boundary perceived from a temporal change, readers are slower and less able to retrieve information prior to the event boundary. This interference of the current event model in retrieving memory from previous events can be seen in personal experiences as well. Radvansky and Copeland (2006) have shown that memory accuracy falls when people walk through doorways.

Studies have also investigated the relationship between event segmentation and long-term memory. The gating mechanism of prediction errors suggests that information perceived during an event boundary will be processed more and subsequently better retrieved from long-term memory (Swallow et al., 2009). Schwan and Garsoffky (2004) found that participants performed better at event recall and recognition tasks after watching movies that preserved event boundaries and omitted non-boundaries compared to movies that preserved non-boundaries and omitted boundaries. Swallow and colleagues (2009) also showed that people had better memory of objects in the background of movies seen during event boundaries compared to those shown during non-boundaries.

Event boundary perception also affects what information is bound together in episodic memory. Ezzyat and Davachi (2011) asked participants to read narrative stories and completed a cued-recall task about two sentences that either had or did not have an event boundary between them. They found that participants' long-term associative memory was worse for information across event boundaries compared to information from within the same event. Functional magnetic resonance imaging data collected during encoding showed that brain activity related to information integration within events had a high correlation with within-event binding. These results illuminate that event segmentation during encoding affects the way that those events are later remembered.

Working Memory Load and Event Segmentation

Despite the abundance of results in line with the assumptions that can be drawn from the event segmentation theory, recent studies have provided evidence suggesting that event boundary perception could occur even in the absence of a prediction error. Heusser and colleagues (2018) found that event segmentation occurs at points of predictable context changes. Similarly, Schapiro and colleagues (2013) showed that events can be inferred based on temporal structure, even when predictability is held constant.

In this study, we tested an alternative theory of event segmentation: that event boundaries are primarily related to changes in the content of working memory. During perception of a continuous stimulus, we must decide how to manage our limited working memory capacity, to ensure that we only maintain relevant information necessary for understanding ongoing actions.

When we infer that the environment has shifted into a new kind of event (whether this shift was predicted or not), we should choose to forget information about the now-irrelevant recent past to avoid interference. We hypothesized that this kind of intentional updating of working memory components could trigger event segmentation in the absence of prediction error. In our first experiment, we created a task that required participants to forget the contents of their working memory at the end of specific segments. Participants were either placed in a "clock" condition, in which the beginning and end of each segment were predictable, or a "no clock" condition, in which the end of each segment was unpredictable, creating a prediction error (an unexpected surprise) about the current segment. We then identified whether participants perceived event boundaries at the points laid out by the segments set within the task by checking the temporal order memory of the images the participants saw. Heusser and colleagues (2018) showed that people were worse at remembering which item was shown first when the pair of items were perceived from two different events compared to when they were perceived from within a single event. We used the same temporal order judgment task that the researchers used in this experiment. Thus, we expected participants would have worse temporal order memory for items from different segments compared to items from the same segment in both the experimental and control conditions, suggesting that event boundaries were perceived when they were actively resetting the items held in their working memory, whether or not they were experiencing an unexpected change.

We also conducted a follow up study that used a set of repeated colored backgrounds instead of clocks by designating a target color that signaled the beginning of a new segment. We compared the temporal order memory accuracy of this experimental condition to a control condition that required participants to perform a 2-back working memory task of identifying whether a given image was repeated from 2 images ago. In this control task, participants were presented with the same visual stimuli as the experimental condition, but the task did not have event structures. We expected to see an event-related memory difference in the background condition but not in the 2-back condition.

Experiment 1

Method

Participants

Eighty-one participants (female = 32, male = 59, M_{age} = 25, age range: 18-52 years) were recruited from an online platform called Prolific. Considering that Prolific is available world-wide, we only recruited subjects who had reported they were fluent in English in order to ensure that participants would be able to understand and follow the instructions. All participants gave informed consent through an online questionnaire and received monetary compensation for their time.

Materials

432 images were selected from the Bank of Standardized Stimuli (Brodeur, Guérard, & Bouras, 2014) and THINGS database (Hebart et al., 2019) to create 2 image stimuli sets, each consisting of 216 items. There was a total of 19 different versions of these two sets, and for each set, images were randomly picked from the pool and arranged to include unrepeated images, images that were repeated from within the current segment, and images that were repeated from a previous segment. Participants were randomly assigned to one of the 19 stimulus sets.

Procedure

Before the main experiment, all participants went through a practice trial of the actual experiment to ensure that they understood the given task in each condition. During the first phase of the experiment, participants were presented with a series of random images along with an image of a clock. Participants were asked to respond whether the presented image was new or repeated from within the current segment. The stimulus set was created to incorporate novel images, images repeated within a single segment, and images repeated across multiple segments.

The beginning of each segment was signaled by the clock arm pointing upwards or at the 12 o'clock position. Each segment consisted of either 5, 6, or 7 images. We randomly distributed the lengths so that there were exactly 2 segments with a given number of images, making up a total of six segments. There were two different versions of the clock images that were shown during this phase of the experiment. In the first condition (no clock condition), the clock image shown was either of an empty circle or the same circle with a clock arm pointing up. Figure 1 shows an example of two segments in the no clock condition. The first segment consists of 5 items. Note that the correct response for the teacup image on the third position is "repeated" since it had been shown previously within that segment, but the teacup image on the seventh position is "new" since the teacup was shown in the previous segment. In the second condition (clock condition), the circle was divided into either 5, 6, or 7 pieces, and the clock arm changed its position along the divided lines. Figure 2 illustrates a 5-item segment followed by a 6-item segment. Since the ticks in the clock image indicates how long the current segment would last, no prediction error would be expected to be perceived at the boundary between different segments. The images were presented for a fixed time of 2.5 seconds (s) during which the participants had to respond, followed by a fixed 2 s inter-trial interval.

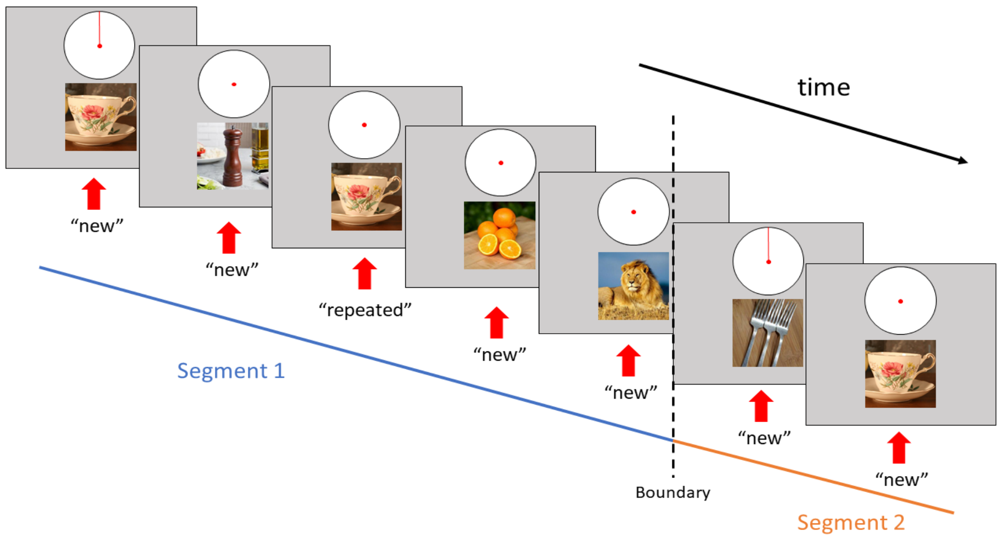


Figure 1. Task sequence example of two 5-item segments in "no clock" condition. Clock image with arm pointing upwards signals beginning of new segment. Note that correct response for teacup on third scene is "repeated" but for same image on seventh scene is "new" since the teacup was not repeated from within the second segment. Prediction error will occur at boundary since beginning of new segment is unexpected.

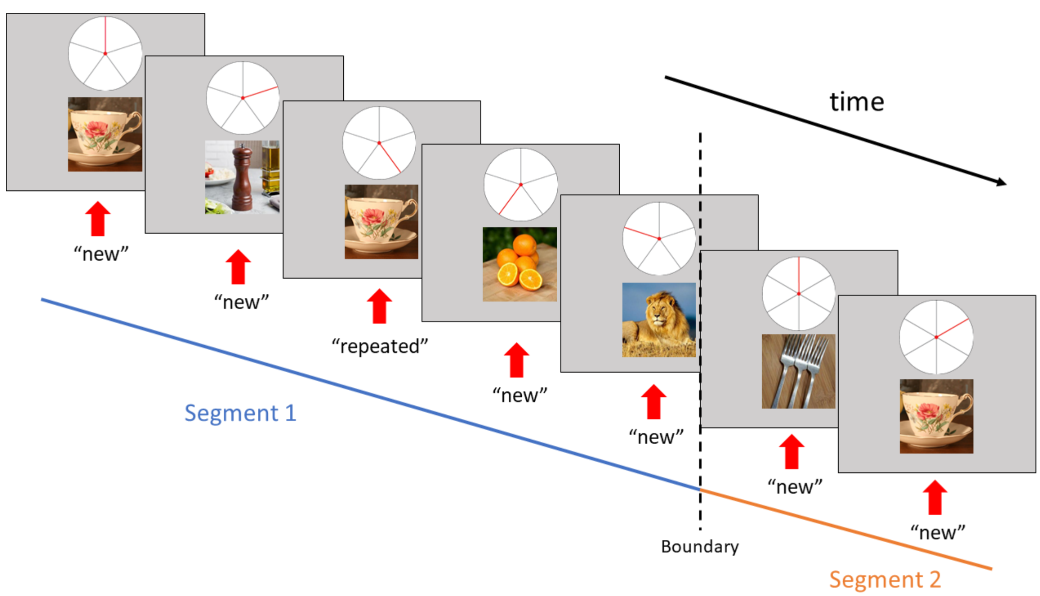


Figure 2. Task sequence example of a 5-item and 6-item segment in "clock" condition. Clock images have ticks that signal how long each segment will last. Prediction error will not occur at boundary since length of each segment is predictable throughout the current segment.

In the second phase of the experiment, we tested the temporal order memory of the presented images. Participants were shown two images from the presentation phase and were asked to respond which of the images they had seen first. For each block, 6 pairs of images were tested: 3 pairs were images from between different segments and the other 3 pairs were from within a single segment. Each pair was chosen from items that were never repeated during the presentation phase and were 3 items apart from each other. The trials were self-paced and advanced once a response was given. There was a fixed .5 s inter-trial interval between test trials.

There was a total of 6 repetitions of the first and second phase of the experiment. Items were never repeated across different repetitions of the phases. All participants went through both conditions. The groups were counterbalanced for order of experimental conditions they completed first.

Results

Working memory task performance

Participants generally performed well in the presentation phase of the experiment with an average accuracy of 95.9% (SD = 3.76%). 6 participants who had performance accuracy below 90% in the task during the first phase were excluded from the final analysis (Appendix A).

Effect of perceptual boundaries on temporal order memory performance

We compared the temporal order memory accuracy of item pairs from within the same event and between different events for both the clock and no clock conditions. Temporal order memory accuracy was used as a measure to identify whether event boundaries were perceived at the beginning and end of segments set out in the presentation phase of the task, as previous research shows that temporal order memory for items from within the same event are more accurate than memory for item pairs from different events (Heusser et al., 2018). For both conditions, permutation tests with 10,000 permutations showed that there was a statistically significant difference between accuracy levels for item pairs from within a single event and pairs from between different events (Figure 3; clock condition: $p < 0.01$, no clock condition: $p < 0.01$). This effect size did not differ significantly between the clock and no clock conditions ($p = 0.21$). These results suggest that participants perceived an event boundary according to the segments signaled by the clock images in the presentation phase in both clock and no clock conditions.

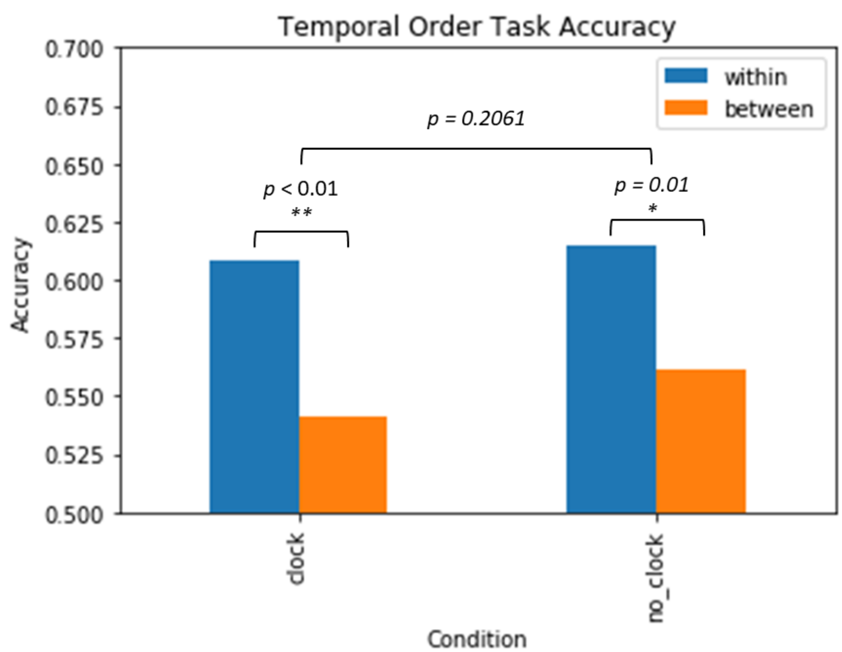


Figure 3. Temporal order accuracy for clock and no clock conditions. Accuracy of items from the same event is significantly greater than accuracy of items from different events for both conditions.

Discussion

Based on the results of this study, it seems that people perceive event boundaries from a series of random objects when updating their working memory load. The effect of event boundaries on temporal order memory was found both when the boundaries for segments were unpredictable and when they were expected from the beginning of the segment. This suggests that event segmentation can occur without a prediction error to trigger an update of the current event model. Our study provides more evidence against the traditional event segmentation theory, and posits that active updating of the contents held in working memory could also drive event boundary perception.

Limitations

While the results from the current experiment provided evidence to support our hypothesis, we identified some limitations that could be addressed through a follow up experiment. First, participants in the clock condition are getting a small amount of new information at the beginning of a new segment as they perceive whether the current segment would be five, six, or seven images long (according to the number of ticks in the presented in the clock). While this set up allowed us to add ambiguity about the length of events to the task, it also could lead to a possible prediction error occurring at these points since this new information is not something that could have been taken into account in advance. We addressed this issue by setting all segments as six images long for the follow up task, providing as little new perceptual information about the task as possible once the main experiment started. Second, there might have been a possible event boundary perception due to the fact that the clock pointing towards 12 o'clock position could have inherent semantic meaning that is related to the end of an event. We used a series of repeated colored backgrounds instead of images of clocks for the second experiment to create a signal that is inherently vague but could be used to clearly indicate segments. Finally, the results from the current experiment provided evidence that a prediction error might not be necessary for event perception, but we wanted to identify whether updating working memory load would be necessary for creating an event-related memory difference for later retrieval. By using the exact same stimuli set and task set up in a control working memory task, we would be able to establish that it was the resetting of working memory that is creating event-related memory effects rather than a variable related to the stimuli or working memory processes itself. Based on these factors, we conducted a second experiment that would address these concerns.

Experiment 2

Methods

Participants

A hundred and twenty-two participants living in the United States (female = 56, male = 60, M_{age} = 32, age range: 18-73 years) were recruited from Prolific. We recruited subjects who had reported they were fluent in English in order to ensure that participants would be able to understand and follow the instructions. Twenty-nine college students from a university in the United States were also recruited. All participants gave informed consent through an online questionnaire and received monetary compensation or course credit for their time.

Materials

The same pool of 432 images that were selected from the Bank of Standardized Stimuli (Brodeur, Guérard, & Bouras, 2014) and THINGS database (Hebart et al., 2019) which was used for Experiment 1 was used to create 4 image stimuli sets, each consisting of 324 items. For each of these 4 sets, items were randomly chosen from the image pool and organized to incorporate the following groups to with different probabilities: 1) items that were completely novel and unrepeated, 2) items that were repeated from within the current segment, 3) items that were repeated from exactly 2 images ago, 4) items that were repeated from a previous segment. The same image stimuli set was used for both the control and experimental conditions.

Procedure

There were two different conditions for this experiment, and participants only completed one of the two conditions. On each trial, an image was presented with a colored background border that was unrelated to the content of the image. The colored backgrounds were picked from a perceptually uniform spectrum and changed along 6 steps from blue to purple (CIELAB). In the experimental condition (background condition), participants learned that the background color changing to blue marks the beginning of a new segment, and were asked to respond whether the presented image was new or repeated from within the current segment they were in. Figure 4 shows an example of two segments in the background condition. This condition was similar to the no prediction error condition from Study 1. The image could be completely novel, repeated from within the same segment, or repeated across multiple segments. Note from Figure 4 that a boundary is marked after the sixth image and the seventh image is considered "new" since it is repeated from a previous segment. Each segment was always 6 items long, making the shift from one segment to the other consistent, observable, and predictable. The images were presented for 2.5 s during which the participants responded, and was followed by a fixed 2 s inter-trial interval.

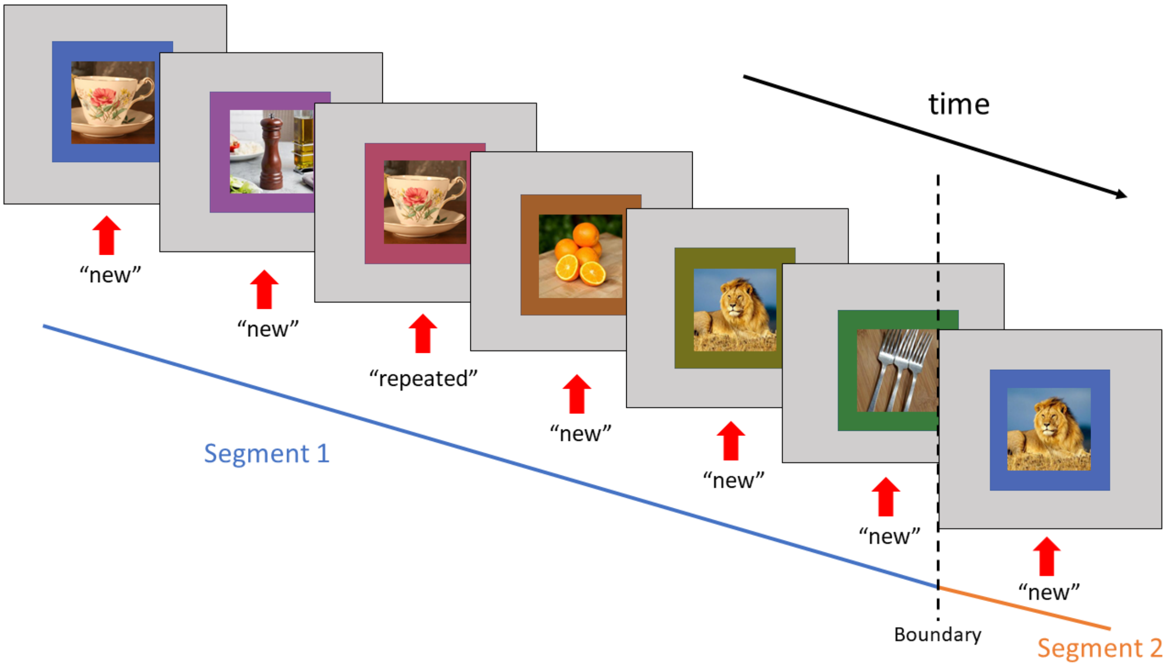


Figure 4. Task sequence example for "background" condition. Background color changes in order of Blue-Purple-Red-Orange-Yellow-Green sequentially and is repeated with Blue indicating the beginning of a new segment. Each segment is 6 images long. Note that correct response for last image seventh scene is "new" since it is the first time the lion image is shown within that particular segment.

We compared the performance of participants that completed the aforementioned task to a control working memory task (2-back condition). In this condition, participants were given the same sequence of stimuli as the background condition task, but were told to respond if they saw an image repeated from exactly 2 images ago. Figure 5 illustrates an example of the 2-back condition task. Note that the correct response for the seventh image is "repeated", unlike in the background condition, since the fifth image that was presented matches it. The participants were instructed that they will see a colored background along with the image, but that this color will not be relevant to the task they would be performing. Since in the 2-back condition, the points in which participants are resetting their working memory load does not coincide with the marked boundaries we are using for analysis, we would expect to not see an event-related difference in temporal order memory measured during the second phase of the experiment.

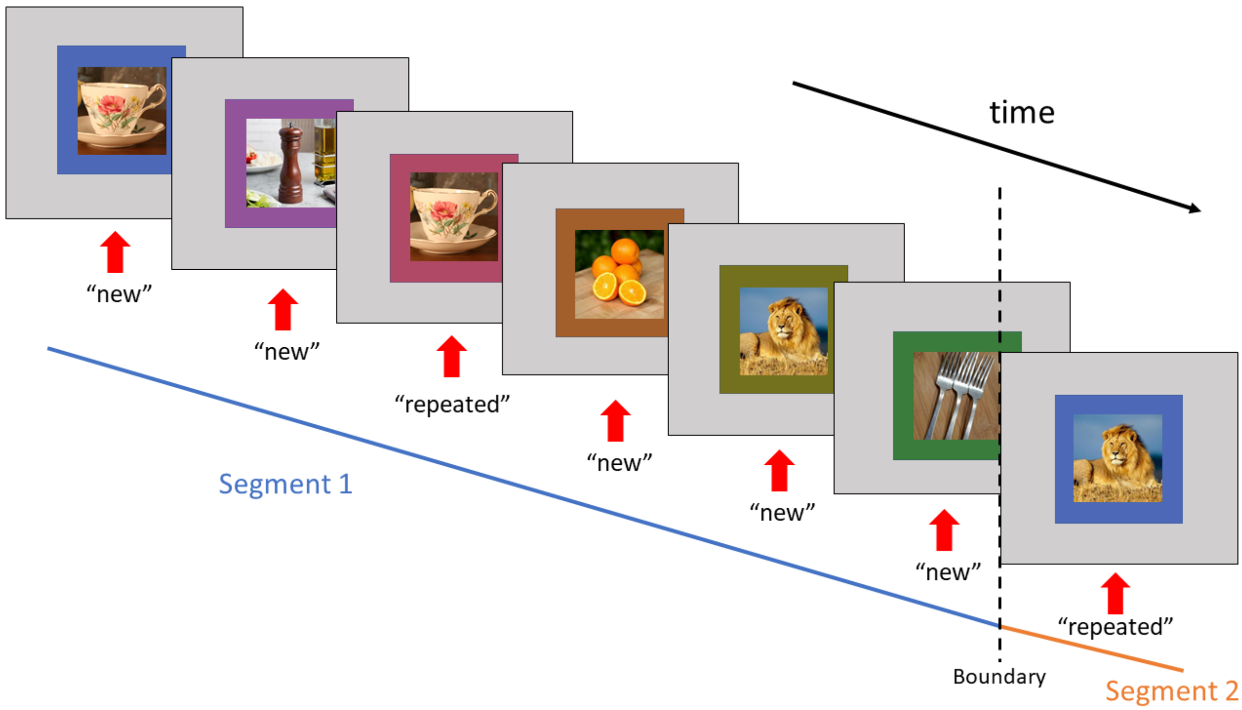


Figure 5. Task sequence example for "2-back" condition. Note that correct response for third and seventh scene is "repeated" since the image is repeated from 2 scenes ago. Participants were instructed that the colored background is not relevant to the task they are asked to complete.

As in Experiment 1, the second phase of Experiment 2 was designed to measure the temporal order memory of items that participants saw during the first phase. Participants in both conditions were shown pairs of images that were presented either in the same segment or in different segments from the first phase. These items were selected from images that were never repeated and the two images had been shown 3 images apart. For each block, 6 pairs of images were tested, 3 pairs from the same segment and 3 pairs from different segments. The trials were self-paced and advanced once a response was given. There was a fixed .5 s inter-trial interval between test trials.

There was a total of 9 repetitions of the first and second phase of the experiment. Items were never repeated across different repetitions of the phases. Each participant went through only one of the two conditions.

Results

Working memory task performance

We have gathered data from a total of 62 participants for the background condition. Most of these participants performed well in the presentation phase of the experiment, with 5 participants performing much worse than 80% accuracy. After excluding data from the 5 individuals, the average accuracy was 92.8% (SD = 4.76%) (Appendix B).

For the 2-back condition, we gathered data from a total of 90 participants. Interestingly, along with some outliers with accuracy below 80%, we also observed a bimodal distribution in presentation phase task accuracy (Figure 6). This suggested that there might be a different factor outside of our manipulation that could be affecting the cognitive processes of the participants, so we performed further analysis of the performance data to identify whether they were completing the task as they were instructed to.

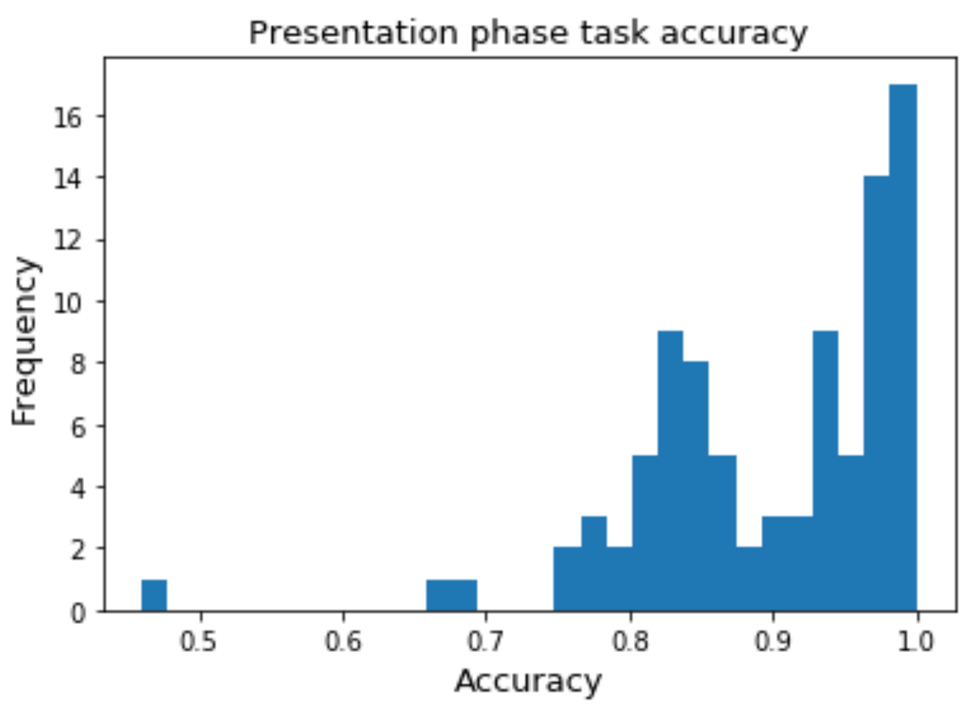


Figure 6. Overall accuracy on working memory task for 2-back condition.

Item-specific working memory task performance by condition

Since we had used a set of stimuli that incorporated images that were repeated from within the segment while not being a 2-back item, images that were repeated from a 2-back but not from within the segment, images that were both repeated as a 2-back and from the same segment, and items that were unrepeated, we were able to analyze the items in which participants were responding wrong to identify whether they were completing the tasks as instructed.

First, we investigated how participants in the background condition were responding to items that were repeated as a 2-back but were not repeated from the same segment. For these images, the correct response would be to identify that they were "new" because even though the images were shown in a previous segment, the task instructions that they received were to reset their working memory at a new segment. As shown in Figure 7, about 45% of the participants from the background condition had less than 50% hit rate for items that were exclusively 2-back. A low accuracy on 2-back only items for this condition could indicate that participants were responding to any image that was repeated, disregarding whether they were from the current segment or not. We excluded the data from the participants that performed worse than at chance level at 50% for this specific category since it would suggest that they were not performing the task that they were instructed to complete. This is a better measure than overall accuracy during the presentation phase of the experiment since there are varying numbers of the category specific items which would allow the participant to show high overall accuracy even if they were not responding in a manner that was required from the specific condition.

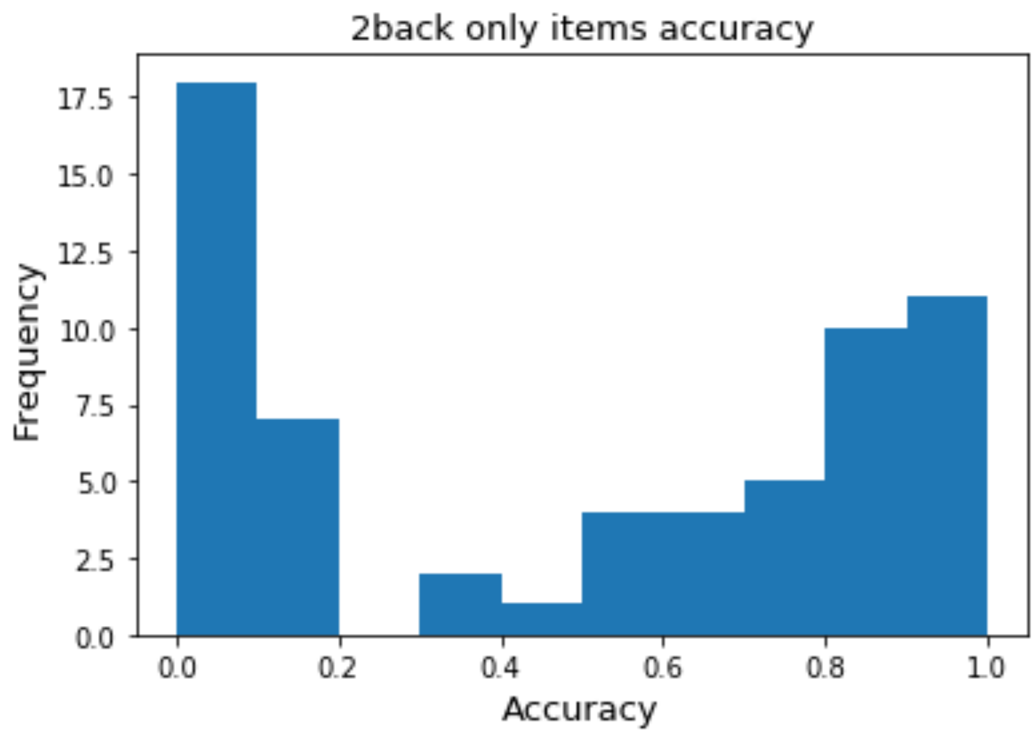


Figure 7. Performance of participants in background condition on items that were repeated from 2 images ago but not from within the current segment.

We also performed a similar analysis with the data from the 2-back condition. For this condition, we looked at the accuracy of items that were repeated from within the same segment but were not repeated from 2 images ago. As shown in Figure 8, there was a very clear divide between participants that were performing well and those that were showing extremely poor performance. A low accuracy score on this category would also indicate that the individual is incorrectly responding that items that were shown before but not from 2 images ago as a "repeated" item. One possible reason for a low score would be if participants were responding to all images, they have previously seen throughout the trial, whether or not it was specifically repeated from 2 image ago. We also excluded data points from participants that had accuracy levels below 50% or chance level from the analysis of the temporal order memory task performance.

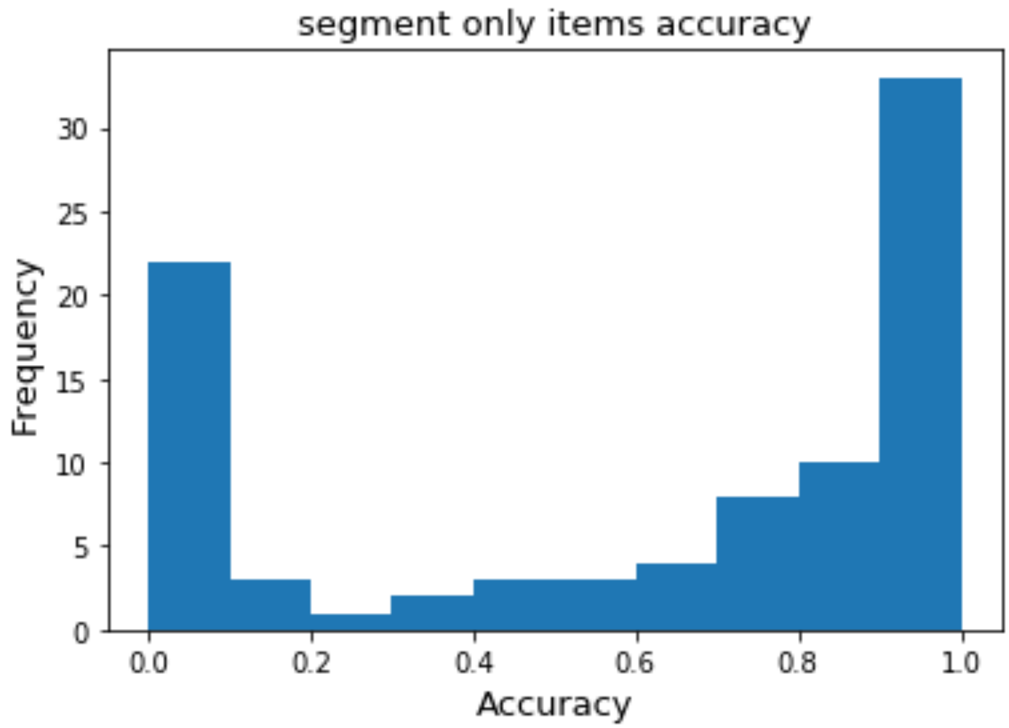


Figure 8. Performance of participants in 2-back condition on items that were repeated from within the current segment but was not repeated from 2 images ago.

Interestingly, the bimodality that was found in the overall accuracy for the 2-back condition was accounted for once we investigated the relationship between accuracy on overall task and on segment only items. Figure 9 shows a scatterplot of overall accuracy and hit rate for images that were repeated from the same segment without being a 2-back repeat. The clusters that are formed suggest that the data points that are on the lower end of the segment only accuracy account for the lower accuracy scores for the overall task.

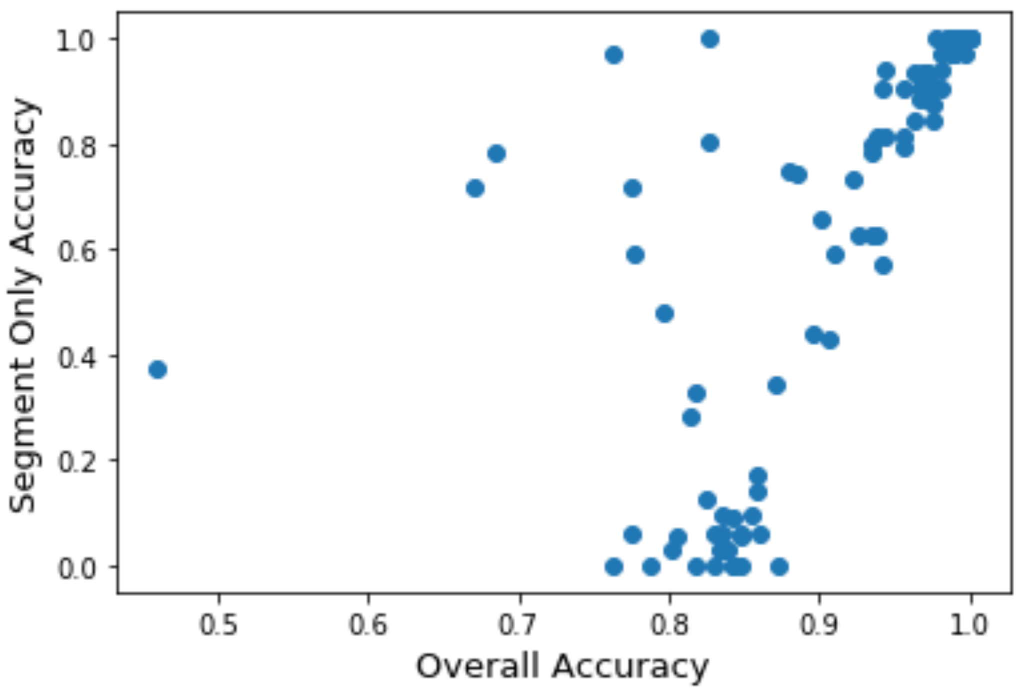


Figure 9. Relationship of overall accuracy and accuracy of items that were repeated from within the current segment but not from 2 images ago for the 2-back condition participants.

Effect of perceptual boundaries on temporal order memory performance

Based on the criteria set above, we excluded data from participants with low accuracy during the working memory task in the first phase of the experiment. Using a similar method as the analysis from Study 1, we ran a permutation test with 10,000 permutations on the temporal order memory task performance for 32 participants from the background condition and 46 participants from the 2-back condition (Figure 10). The results showed that there was a statistically significant difference between accuracy levels for item pairs from within a single event and pairs from between different events for the background condition ($p < 0.01$). However, this effect was not found for the 2-back condition ($p = 0.07$). More importantly the difference between the within and between accuracies for the background condition and 2-back condition showed statistically significant differences, indicating that the event boundary related memory effect was only present in the background condition ($p < 0.05$). This suggests that an event boundary was perceived at the segments delineated by predictable and consistent cues through the active updating of working memory.

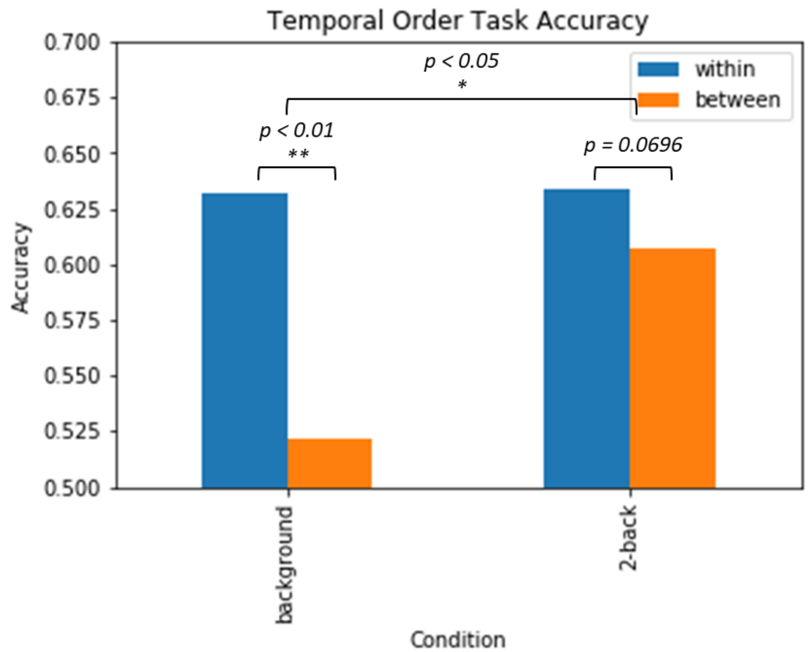


Figure 10. Temporal order accuracy for background and 2-back conditions. Accuracy of items from the same event is significantly greater than accuracy of items from different events for the background condition but the same effect is not shown in the 2-back condition.

Discussion

These results from Study 2 replicated our findings from Study 1 in that event boundaries can be perceived in the absence of a prediction error from a random string of common objects through the resetting of working memory load. This is evident from the comparison with a scenario in which participants were given the same task visual stimuli but were not given structural points to update their working memory and were rather expected to continuously maintain the last two items they have seen, as this control condition showed to not have the same effect of perceived boundaries. Since we used the same task structure and stimuli for both conditions, we can rule out that the boundary effect found for the background condition group is being created from any other perceptual stimuli that the participants are receiving.

General Discussion

Event segmentation is a natural cognitive process that we engage in to better understand and perceive the world that we live in. Previous studies have shown support for the theory that event boundaries are marked when the individual's prediction about the future goes against the perceptual input gathered and leads to a prediction error (Zacks et al., 2007). However, the results from our two experiments support the hypothesis that updating the contents of one's working memory can trigger the perception of an event boundary, without any prediction error. In our first experiment, we created a new task in which participants had to update their working memory of random images at the beginning of segments that were either clearly predictable or not. We found that participants in both conditions showed significantly better accuracy for temporal order memory of items that were from the same segment compared to those from different segments, consistent with the boundary-related effects on memory found from previous research. This suggests that resetting working memory load at the given segments in the absence of a prediction error led to event boundary perception as well as the condition in which prediction error occurred. Our second experiment expands on this finding while using consistent changes in background color to make the beginning of a new segment predictable. The predictable background condition led to higher temporal order memory accuracy for within segment items compared to between segment items while a similar working memory condition did not show the same effect. Since we minimized any possible prediction error occurring in the second experiment, along with using the same stimuli and task set up as a control task, it added additional evidence to our hypothesis that the active reset of working memory at specific locations drives event segmentation rather than other perceptual aspects such as the stimuli or prediction error.

These results highlight the importance of understanding the relationship between working memory and event perception. Event segmentation theory posits that event models that are used to create predictions are held in working memory (Zacks et al., 2007). Our findings add to this theory by providing evidence that prediction error might not be the only factor that affects the updating of the current event model, but rather, the act of resetting the contents of working memory is the driving force that marks the beginning and end of specific events. This could also aid in providing explanations of why event segmentation could also occur during predictable and expected changes in perceptual input, a situation that seems to be unaccounted for from the conclusions of event segmentation theory.

The current study could be improved to provide further evidence for our hypothesis that working memory reset drives event boundary perception. First, we had to exclude a number of participants in our second experiment due to noncompliance. Part of this could have been due to some misunderstanding that was difficult to resolve due to the fact that participants were completing the task via an online platform, creating a burden to reach out for clarification or questions compared to experiments conducted in the lab. For future data collection, it would be crucial to add more detailed instructions and a step to check compliance within the task before moving on to the main task. Also, unlike other event perception research projects, we focused on using a series of random images rather than complex narrative stimuli or film clips to prompt segmentation. The perception of a series of images might be quite different from the cognitive processes that we engage in throughout our everyday life. Thus, creating a method of testing our working memory hypothesis in the context of more naturalistic and complex stimuli would add more insight to the main theory of event segmentation.

The current study investigates working memory reset as a factor that could contribute to event boundary perception, even in the absence of a prediction error. Growing evidence that prediction error might not be necessary for event segmentation suggests that it might be necessary to update the way that we think about the processes involved in the perception of event boundaries.

References

Ben-Yakov, A., & Henson, R. N. (2018). The Hippocampal Film Editor: Sensitivity and Specificity to Event Boundaries in Continuous Experience. *Journal of Neuroscience*, *38*(47), 10057–10068. https://doi.org/10.1523/JNEUROSCI.0524-18.2018

Brodeur, M. B, Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. PLoS ONE *9*(9): e106953. https://doi.org/10.1371/journal.pone.0106953

Ezzyat, Y., & Davachi, L. (2011). What Constitutes an Episode in Episodic Memory? *Psychological Science*, *22*(2), 243–252. https://doi.org/10.1177/0956797610393742

Hard, B. M., Tversky, B., & Lang, D. (2006). Making sense of abstract events: Building event schemas. *Memory & Cognition, 34*, 1221-1235.

Hebart, M. N., Dickter, A. H., Kidder, A., Kwok, W. Y., Corriveau, A., Van Wicklin, C., & Baker, C. I. (2019). THINGS: A database of 1,854 object concepts and more than 26,000 naturalistic object images. PLoS ONE 14(10): e0223792. https://doi.org/10.1371/journal.pone.0223792

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *44*(7), 1075–1090. https://doi.org/10.1037/xlm0000503

Magliano, J. P., Miller, J., & Zwaan, R. A. (2001). Indexing space and time in film understanding. *Applied Cognitive Psychology*, *15*(5), 533–545. https://doi.org/10.1002/acp.724

Magliano, J. P., Radvansky, G. A., Forsythe, J. C., & Copeland, D. E. (2014). Event segmentation during first-person continuous events. *Journal of Cognitive Psychology*, *26*(6), 649–661. https://doi.org/10.1080/20445911.2014.930042

Radvansky, G. A., & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition*, *34*(5), 1150–1156. https://doi.org/10.3758/BF03193261

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. - PsycNET. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 210–216.

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*(4), 486–492. https://doi.org/10.1038/nn.3331

Schwan, S., & Garsoffky, B. (2004). The cognitive representation of filmic event summaries. *Applied Cognitive Psychology*, *18*(1), 37–55. https://doi.org/10.1002/acp.940

Speer, N. K., & Zacks, J. M. (2005). Temporal changes as event boundaries: Processing and memory consequences of narrative time shifts. *Journal of Memory and Language*, *53*(1), 125–140. https://doi.org/10.1016/j.jml.2005.02.009

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human Brain Activity Time-Locked to Narrative Event Boundaries. *Psychological Science*, *18*(5), 449–455. https://doi.org/10.1111/j.1467-9280.2007.01920.x

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*(2), 236. https://doi.org/10.1037/a0015631

Zacks, J. M. (2010). *How We Organize Our Experience into Events* [Data set]. American Psychological Association. https://doi.org/10.1037/e554772011-002

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*, 979-1008.

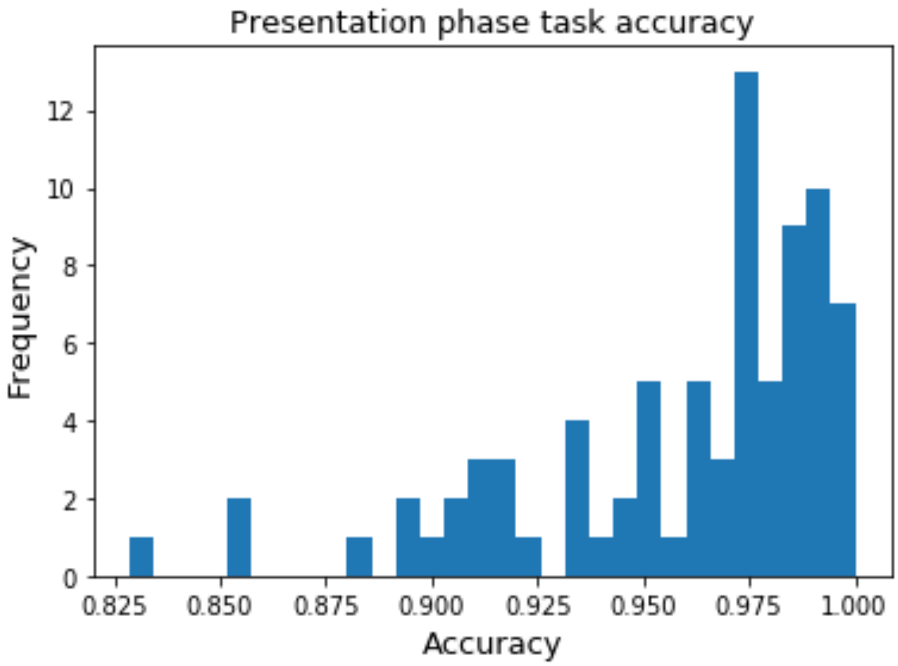
Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries | Nature Neuroscience. *Nature Neuroscience*, *4*, 651–655.

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction Error Associated with the Perceptual Segmentation of Naturalistic Events. *Journal of Cognitive Neuroscience*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078

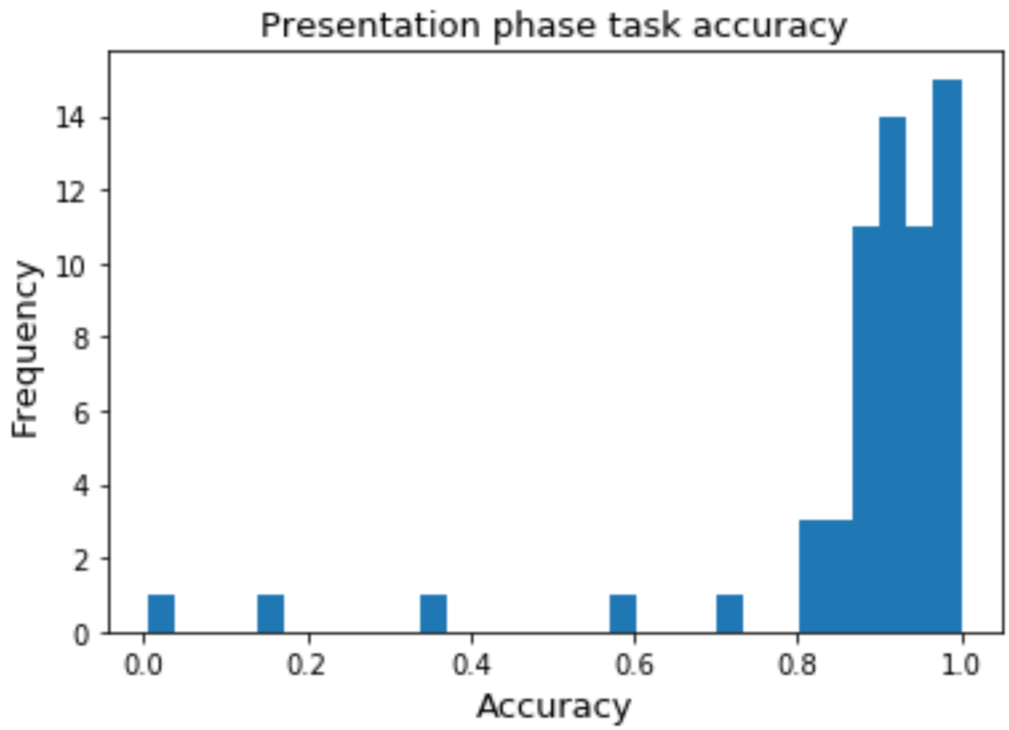
Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305

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

Appendix A


Presentation phase task accuracy
Frequency
0
2
4
6
8
10
12
0.825
0.850
0.875
0.900
0.925
0.950
0.975
1.000
Accuracy

Appendix B


Presentation phase task accuracy
Frequency
Accuracy
0
2
4
6
8
10
12
14
0.0
0.2
0.4
0.6
0.8
1.0