Movies & Memory:
How Film Editing Can Impact Episodic Memory Formation

A thesis presented by

Jake Schwencke

To
The Faculty of the Committee on Degrees in Neuroscience
In partial fulfillment of the requirements
For the degree with Honors
Of Bachelor of Arts
And Certificate in Mind, Brain, & Behavior

Harvard University
Cambridge, Massachusetts

March 12, 2021

Neuroscience Concentration

Division of Life Sciences

Harvard University

The Harvard College Honor Code

Members of the Harvard College community commit themselves to producing academic work of integrity – that is, work that adheres to the scholarly and intellectual standards of accurate attribution of sources, appropriate collection and use of data, and transparent acknowledgement of the contribution of others to their ideas, discoveries, interpretations, and conclusions. Cheating on exams or problem sets, plagiarizing or misrepresenting the ideas or language of someone else as one's own, falsifying data, or any other instance of academic dishonesty violates the standards of our community, as well as the standards of the wider world of learning and affairs.

Signature:________________________Jake Schwencke_______________________________

Acknowledgements

If there were a section acknowledging all the reasons this research should not have happened, it would be twice the length of this thesis. Instead, I'd like to acknowledge the few key players, advisers, confidants, and friends that made it possible for a storyteller to dive deep into the throws of hard science-- to play and explore at the intersection of movies and the mind.

Firstly, I cannot find enough words to express my gratitude for my postdoctoral partner-in-crime, Jie Zheng who guided me through this process. Through over a year of weekly meetings, she embraced my ideas with open arms, providing further research and consistent advice in order for us to craft a legitimate research project. I could not have come close to the finish line without her support, knowledge, and enthusiasm for these ideas.

I am also indebted to Gabriel Kreiman for welcoming me into his lab based on a proposal that had raised eyebrows on both the scientific and artistic side. A computational scientific research lab and a non-computational filmmaker does not seem like the most obvious match, but his willingness to connect my interests with Jie's ongoing research was crucial in allowing my ideas to come to fruition.

Speaking of computation, I must also thank my in-house coder and incredibly generous roommate, Daniel Nguyen, for spending countless hours with me as I coded my results and analysis. As if joint-concentrating wasn't enough, Daniel was a fantastic, patient professor of my pseudo-secondary in Computer Science for Theses.

I must also thank my thesis fairies and professional pep-talkers Jenna Lang, Sydney Robinson, Colin Gray-Hoehn, and Bobby Malley for putting up with years of my ranting about how movies target the mind. They always reminded me I'm only half crazy.

A huge thank you goes out to my mom and dad for being the subjects of our stimulus set, filmed at the height of the pandemic. Even now I'm not so sure they understand what is happening in this research, but they have provided unwavering support nonetheless.

Finally, I'd like to thank my neuroscience adviser, Ryan Draft, for being the first person at Harvard to tell me this joint-concentration was possible. He took my dreams of a course of study that used movies to think about the mind seriously, and mapped out a path to make it happen. It is thanks to these individuals that I am able to present the following thesis.

LIST OF CONTRIBUTIONS

Building off of prior research and postdoctoral researcher Jie Zheng's previous work on neural responses to cognitive boundaries, this project was conceived by myself, Jake Schwencke, and Jie Zheng. We, in collaboration with Gabriel Kreiman, contributed ideas for experimental design and analysis. I created the filmed stimulus set while Suneil K. Kalia, Taufik A. Valiante and Adam N. Mamelak managed the patients and surgeries. Jie Zheng, Andrea G.P. Schjetnan and Mar Yebra collected data. Jie Zheng and I jointly wrote code to analyze the behavioral and neural data. I performed the analyses of the results with assistance and guidance from Jie Zheng.

ABSTRACT

While we experience life as a continuous stream of input, memories are organized as discrete events. Event Segmentation Theory states that cognitive boundaries segment and structure memory, but the underlying mechanism for this process remains unclear. We hypothesize shifts in neural response within the medial temporal lobe reflect event segmentation and influence subsequent memory retrieval. To test this, we recorded activity from single neurons in patients with drug-resistant epilepsy during the viewing of movie clips. Patients were tasked with recalling the clips and order of events within them. We used “cuts” as visual cues for event segmentation and evaluated how the manipulation of cut locations can influence the temporal structure of episodic memory. Behaviorally, subjects tended to have improved memories of the temporal order of events from clips where cuts coincided with event boundaries. Neurally, we found single-units that responded significantly to abstract cognitive boundaries between events. Some of these neurons responded to a specific boundary in the clip, presumably maintaining temporal structure of events. This increased neuronal firing in the encoding associated with cuts that coincide with event boundaries may be a neural explanation for the improved temporal discrimination accuracy as well as an indication of encoding for temporal structure. Overall, these results reveal neural signatures associated with event segmentation as well as the impact an editor can have on an audience’s memory in the making of a film.

AUTHOR'S NOTE

It had been hours of me floating-- freezing. I noticed a light in the distance and tried to rouse him but he wouldn't wake. Others surrounding me, who once thrashed about in the water, now bobbed in their life vests, still. The words, "WRITTEN AND DIRECTED BY JAMES CAMERON" appeared on the screen, and the lights came up, revealing my face soaked with tears. I had spent three hours transported from my home to the world's largest steam vessel. I flinched as water flooded the mere idea of a hallway. I mourned the forbidden love of a series of pixels, watchingthe 1997 film, *Titanic*.

I think of storytelling as a phenomenon of the cognitiveexperience, and I admit to crying alongside Kate Winslet to exemplify that. Film and television have a profound effect on us, eliciting emotions that resonate often longer than lived experiences. We consume fictional narratives to elicit real emotions, often dismissing these stories as momentary entertainment instead of appreciating them as vessels for communication and empathy. I am constantly amazed by how skillfully movies capture my attention with contrived narratives, and how vividly I can recall the stories within them. It is this passion for movies, filmmaking, and acting that has led to my deep interest in studying the brain.

This thesis explores the intersection of filmmaking and neuroscience, specifically how much influence a filmmaker can have over an audience's memory. After all, the structure of a film is carefully edited to be understandable, impactful, and memorable. To successfully string together a collage of discontinuous clips into a coherent and convincing narrative, one must know, or at least intuit, a great deal about cognition. Simply put, a director without a sense for how actions and events are perceived by their audience will never make a blockbuster. I hope for this thesis to provide a sense of the possibilities that lie at the intersection of movies and the mind, and in this case, specifically, memory.

WHY DO MOVIES WORK?

We react to movies and television in the same way that we experience the real world. In fact, it has been shown that the same mental responses one has when watching a film parallel the responses made when participating in real world events (Bezdek et. al, 2013). But how could a series of pixels projected onto a screen elicit the same mental responses brought on by the physical world? In short, we did not evolve to watch movies. Our visual processing streams have developed over thousands of years, with only the most recent 200 years including the moving image. We evolved to experience reality; to respond appropriately to stimuli with the purposes of increasing survival (Zacks, 2015). As a result, when a car veers toward the screen or a tiger viciously leaps toward the camera, we flinch because our visual processing streams are wired to react appropriately.

While evolution has hard wired our brains to instinctively respond to the visual stimuli we encounter, our everyday visual input comes in the form of a steady, continuous stream of action (Murch, 1992). Never have we glanced out the window of a three-story building and instantly adjusted our field of view to a close-up of the pedestrians below. For a better view, we have to squint to make out their faces or drag ourselves down the stairs to see them clearly. Yet, movies constantly switch camera angles, to allow the spectator a perfect view of the action taking place. These edits (when the camera's point-of-view switches to a new angle), cause huge transient visual changes, often causing every pixel of the screen to change in an instant.

This phenomenon is distinct from our evolved visual experience, yet we so readily accept these gigantic discrepancies. How then, can a series of shots, depicting fake events, taken from different angles and at different times come together to so effectively convince me that the Titanic is sinking in front of my very eyes? Where does this ability to effortlessly integrate these displays come from? This marks the crossroads between filmmaking and cognitive neuroscience.

We may not have evolved to watch movies, but movies have evolved to respond to the needs of our visual processing streams. In particular, filmmakers make great use of our curious and predictive nature. Our brains are constantly grappling to make sense of the world, asking questions about what we experience. We instinctively make predictions about ongoing action, and search for the correct answer (Zacks, 2015). By editing films in a way that answers these ongoing visual questions, the filmmakers satisfy our visual processing streams. Thus, we can attend to the higher-level meanings within the story (Hasson, 2008). The meaning of the story then occupies enough attentional resources to gloss over the abrupt visual change that provided us with those visual answers. For example, when a character looks off screen, the following shot will almost always be a close-up of the object the character is focused on. Similarly, if a character reaches for a glass, the next angle will show the hand arriving at the cup. These "Point-of-View" and "Gaze Match" edits answer questions about the object the character is interacting with. The success of filmmakers' intuitions to employ point-of-view/gaze match edits sparked research into where individuals attend during action. It was discovered that people in real situations reliably look to the object that is being reached for, and this has been found to be true in both actors and observers (Eisenberg et al. 2018, Flanagan & Johansson 2003, Hayhoe & Ballard 2005). Once again, films are taking advantage of individuals' perceptual proclivities without even realizing it.

Figure 1: Shot/Reverse Shot From Shawshank Redemption (1994). *Another common editing style called the Shot/Reverse Shot is often employed during conversations. The shots are meant to take the perspective of thenon-speaker, looking over their shoulder, forming a proper view of thespeaker. This mirrors where an observer would attend as they look back and forth at who is speaking.*

These types of editing rules or *Continuity Editing Rules*, are so effective, they bring rise to a remarkable phenomenon referred to as *Edit Blindness*. Film edits can go completely unnoticed, even when asked to attend to them (Cutting & Candan 2013, Shimamura et. al, 2015). Edits adhering to continuity editing rules result in greater edit blindness than those that do not, with up to a quarter of within scene edits deemed "invisible". This increases to a third when the edits occur across motion, known as a "Match-Action" cut (Smith, 2012). By cutting across a shift in attention to a new speaker or during a sudden onset of motion, the editor can increase the likelihood of failure to notice an edit as a result of forced attention to information guided by the visual narrative. Continuity Editing Rules like "Match-action" cuts and "Shot/Reverse Shots" allow for the seamless stitching of disparate clips to form a sequence that feels natural and intuitive to the viewer, and Hollywood has coined these techniques to tell effective visual stories (Murch, 1992).

Another theory surrounding how filmmakers so effectively transport viewers into their fictional worlds is through an edit's exogenous control over a spectator's saccades. Happening 2-4 times a second, saccades are rapid micro-movements our eyes make to survey the visual field. We tend to saccade to faces and relevant objects with more frequency (Zacks, 2015). The movement itself takes about 50ms, but it also causes diminished visibility for up to 150ms before, during, and after these saccades (Cutting, 2013). When an edit occurs, viewers saccade to the center of the screen as a reflexive response, to begin surveying the new stimuli. Editors mostly place the attentional focus at the center of frames or on faces, and are therefore redirecting our attention to the target at hand with every edit. In this way, films develop sly control of our attention through the building blocks of their own structure (Cutting & Candan, 2013). In fact, Tang et al (2016) showed just how consistent eye movements are when watching a movie. Coupled with the diminished visibility induced by saccades themselves, editors can control where we look and odds are we won't even take notice.

The goal of continuity editing is to combine disparate camera angles into coherent, seamless narratives. This process requires two distinct achievements: *Conceptual Integration*, when the audience concludes that two shots are representing the same event, and *Perceptual Continuity*, meaning that the characteristics of both shots match enough so as to not induce any distractions that lessen the involvement in the story (Levin & Baker, 2017). In this way, filmmakers can tell a spatially and temporally coherent narrative that can engross the viewer enough to transport them into the world of the movie. Once strong conceptual integration and perceptual continuity are achieved, the film's narrative, to the viewer, can appear real. We experience the world through the lens of the filmmaker and choices of theeditor. I discuss later how the edits themselves, in the context of what we have come to know about event perception, can shape the way we internalize the stories we watch.

PERCEPTION AND MEMORY

By considering the nature of our brain's authentic response to visual stimuli and Hollywood's ever present goal to tell increasingly engaging stories, I realized there may be an opportunity for research that benefits the pursuits of both art and science. How we perceive and recall events is of interest to both a psychological scientist and an eager storyteller. After some more digging, it soon appeared I was not the first to have this idea. Hugo Munsterberg, a German-American psychologist and Harvard professor wrote *"The Photoplay: A Psychological Study"* in 1916, about a machine that displayed a series of images that appeared to move when spun. He argued that cinema was a natural domain for psychological study. He referred to it as 'mental play', and that the narratives filmmakers created could reveal mental processes like the structure of thought itself, as well as insights into attention, memory, imagination and emotion (Munsterberg, 1916). His ideas were widely ignored for 50 years, but he was an early proponent for taking a cognitive approach to movies. Since then, researchers have begun to explore cinema's potential in helping to understand event perception and memory, as well as helped to find more engaging storytelling techniques through scientific research.

In life, events are ephemeral. In film, these events are able to be manipulated and replicated. The scene of a bustling city block is marvelously chaotic, but on the whole structured, regular, and therefore predictable. But what is the underlying process behind how we intake and process these events? To gain insight into how we perceive ongoing events, Darren Newston and colleagues presented participants with unedited video footage. Newston asked them to press a button every time they felt they observed a natural boundary between events. They discovered that observers could identify meaningful boundaries between events with high intraindividual reliability and interindividual consistency (Massad, Hubbard, & Newtson, 1979, Newtson 1973, Newtson, Engquist & Bois 1977).

This idea of *Event Segmentation*, or dividing a continuum of action into parts, was later characterized by a number of definable characteristics. Not only is there robust intersubjective agreement about the locations of these boundaries, but segmentations are reliably induced by an increase in stimulus change. (Radvansky & Zacks, 2014). A neuro-imaging study involving passive viewing of everyday activities found that when the onset of a new event caused the detection of a predictionerror, a large network of regions in the brain would become active, time-locked to these event boundaries (Zacks, 2001; Zacks, 2010). Universal segmentation was found to be true in both first- and third-person studies, making it applicable for movie watching as well as real life (Hemeren & Thrill 2011, Tauzin 2015, Magliano et al. 2014, Swallow et al, 2018). As a filmmaker, this struck me as particularly intriguing: our visual processing streams intuitively divide life into segments, like scenes of a movie. A coincidence? We will come back to this point later.

Event perception has been found to be both predictive and inferential. Predictive, in the sense that we have a keen eye for identifying goals. Determining an actor's goals allows for our brains to predict and become better prepared for the action that will take place next. From infancy through adulthood, we attend to goal-relevant aspects more than others (Woodward 2009, Papafragou 2010). Even more interestingly, when goal-relevant aspects of movie clips are changed, these aspects are noticed more than altered aspects that do not associate with the action's goal (Loucks & Pechley, 2016). Perception is also inferential, in that we call upon prior knowledge often to predict upcoming events. *Event Schemas* or "scripts" hold knowledge about the structure of events, so relationships between objects can be readily inferred (Abelson 1981). The classic example includes when the lights go down at a concert. Most individuals' concert schemas will infer from past experience that the lights dimming means the next event will be the performers taking the stage. They respond predictively and appropriately by attending to the stage (Zacks, 2020). When monitoring saccades for films placed in a rich context, more systematic eye movements are observed, as context provides a richer event schema leading to more accurate predictions (Loschkey et al 2015).

By attending to goals and making inferences from prior knowledge, we segment ongoing action and use these segmentations to update event models. *Working Event Models* are spatiotemporal frameworks that are theorized to be the mental representation of ongoing perception of an event. These mental models are both a perceptual representation of the current unfolding activity and a working memory representation, as information about objects and relationships within the event are maintained throughout perception. (Radvansky & Zacks, 2014).

In sum, we segment ongoing activity in order to create and maintain event models that in combination with previous knowledge and goal attention helps to make predictions about the ongoing events. It is when these predictions begin to differ from what is being perceived that we mark the onset of a new event, according to *Event Segmentation Theory* (Magliano & Zacks, 2011). Accurate predictions become less likely when approaching event boundaries, and these boundaries correspond with fMRI activity in the midbrain structures, which are associated with signaling prediction error (Huff et. al. 2014, Zacks et al. 2011). Thus, the cycle continues.

These mechanics behind how we perceive the world are thought to be the beginnings of how we process and choose what details of our experience crystalize into memory. Our brains are constantly inundated with the many details that make up everyday life. Only a fraction of that information ever consolidates into retrievable memories. The exact process of selection and interpretation of this information remains unknown. However, it is some form of this subjective filtering process which becomes the basis by which the brain can form a coherent narrative of episodic memory. Event Segmentation Theory provides some insight into this process, as event boundaries are consistently remembered disproportionately well. Viewers remember the moments of greatest feature change like shifts in location (like through doorways), changes in time (flashbacks or across scenes), breaks between situational comedy episodes, etc (Perrijohn et. al 2016, Radvansky et. al 1998, Huff et. al, 2014). This may be an explanation for the success behind the art of the "plot twist": A huge spike in prediction error as the narrative gets turned on its head always sticks with us more than the ending we would have seen coming. It is moments of change and prediction error that tend to solidify into retrievable memories. This, therefore, leads to the idea that 1.) Event segmentation unitizes ongoing activity as we perceive and respond to it, and 2.) This unitization becomes instrumental in how we select and organize what details form a coherent narrative of episodic memory.

NEURAL REPRESENTATIONS OF EVENT BOUNDARIES

The road to taking the cognitive elements of event segmentation theory and investigating it at the neuronal level is challenging, but begins with the rodent neurophysiology and human neuroimaging studies that support segmentation of space. There is reason to believe that the segmentation of experienced events share key mechanisms with that of physical space, as movement through space is closely related to movement through time.

The processing of route development in spatial navigation is well studied, and places particular importance on moments of greatest change: turns and edges. Observing rodents in enclosures, boundary cells found in the hippocampal subiculum fire at edges (Lever 2009; Stewart, 2014; Bird, 2010). Additionally, grid cells in the Entorhinal cortex provide hexagonal symmetrical firing fields that tile the environment (Hafting, 2005). As rodents cross over to new compartments, hippocampal cell firing rates spike. (Spiers, 2015). Interestingly, this corresponds well with individuals' tendency to remember boundaries such as walking through doorways, and representations of flashbacks, as mentioned earlier. This is thought to be a potential behavioral consequence for this spatial remapping (Radvansky et. al 1998).

The hippocampus is a strong substrate for extracting timepoints of relevance, episodic memory formation, and is known for prediction of upcoming states (Tulving and Markowitsch 1998, Eichenbaum 2004, Buzsaki and Moser 2013). In Daw 2005, trends in firing patterns changed as routes became more familiar. As the mouse became familiar with the turns of a specific route, "the likelihood of expectancyin activity decreased for the unrewarded arm in a specific turn" (Daw, 2005). This suggests an increase in prediction signaling for the upcoming changes, similar to the cognitive "Event Schemas" that are thought to help inform one's predictions of a given event. In addition, the posterior medial cortex shows trends in activity that come before peaks in hippocampal activity, which reflects more upon the timescales by which event structure is coded (Brunec, 2018). Lastly, studies using fMRI discovered that objects placed at turns within a given route (moments of greatest change) saw greater activation than objects placed at non-turn points (Janzen, 2004; Janzen, 2007).

In this way, the physical edges and turns of spatial navigation seem to exhibit similar importance to the boundaries between events described in spatiotemporal contexts. They seem to share underlying mechanisms of learning information about the event (whether it be a route or event-specific), and make predictions by extracting regularities gleaned from prior experiences (Brunec, 2018). Other studies have also confirmed robust peaks in hippocampal activity at the onset of transient change between events during video viewing (Ben-Yakov, 2011; Ben-Yakov, 2013). The strength of this hippocampal response, as well as triggered shifts in posterior cortical regions like the posterior cingulate and angular gyrus, can be related to the subsequent memory for these complex video clips (Baldassano, 2017; Chen, 2017). Upon recall, portions of cortex in addition to hippocampus were activated, suggesting a cortical hierarchy that detects not only boundaries, but triggers recall of information from a longer narrative to enrich the current working event model (Brunec, 2018). This similar prioritization toward boundaries strengthens the argument that boundaries are key information for memory formation.

FILM IN THE CONTEXT OF PERCEPTION AND MEMORY

As our minds unitize ongoing events and use these boundaries to organize what we come to remember later, filmmakers unitize the narrative through edits, and use these boundaries, or scenes, to organize their stories. Film editing seemingly mirrors the way we process and filter the world into digestible units. This may be a coincidence, or it may be filmmakers' intuition for cognition.

Overtime films have included more and more edits, instead of long continuous scenes that often populated older films. Some of the first motion pictures were one continuous clip, and slowly these clips were placed next to each other to tell more and more complex stories. Overtime, the storytelling capabilities of the medium have been taken advantage of, and these stories consist of rapid edits thfat create a narrative arc determined by the filmmaker.

Films, therefore, contain two distinct structures that operate in parallel: The natural structure of the depicted events, which is segmented by viewers and instinctively unitized and remembered at boundaries. The other is the edited structure, segmented by the changes of camera angle, which work hard to satisfy the visual processing stream to go unnoticed and also unitize the film into scenes to form a narrative structure. This prompts the question: How might we be able to edit movies in a way that, without disrupting view integration, harnesses the perceptual mechanism that facilitates episodic memory formation? Can we optimize editing in a way that makes films *more* memorable due to strategic placement of edits in accordance with event boundaries? Can the way in which we edit a film narrative influence the memorability or engagement with its events? In this way, we can use neuroscientific research to probe how we form memories and learn about ways to make more engaging, memorable movies.

This has sparked a debate among researchers. The power of editing and its influence on event segmentation and subsequent memory has been a debate due to mixed empirical support. Carroll and Bever (1976) and Kraft (1986) both found that memorability only depends on the activity of the depicted events, and that the presence/absence of cuts does not matter. Conversely, Schwann, Hesse, and Garsoffky (1998) found that the frequency of segmentations after cuts was significantly greater. In our own group's study at the Kreiman Lab, subjects tended to remember frames directly after cuts compared to in the middle of clips (Zheng, 2019). This information leads us to believe that cuts (unpredictable transient changes) can increase error in prediction, therefore facilitating the organization of information and assisting in event segmentation. Another study indicated that placing cuts at natural event boundaries can increase memory recall, mitigating differences from previous studies to attempt settling this debate (Schwann 2000).

Four possible syntactical purposes of cuts are detailed in the table below. These studies present participants with versions of films that either contain edits at universal breakpoints (the boundary between the end of one event and the start of an adjacent event) or edits at non-breakpoints (the middle of an event). The results of Schwann, 2000 largely provided support for the Irrelevance hypothesis. However, an observed increase in recall after watching the Cuts-At-Breakpoints version of a film compared to other versions supported the Accentuation hypothesis, in that accompanying an event boundary with an edit can enhance memory of that moment. Interestingly, the Cuts-At-Nonbreakpoints version saw a negative correlation between the number of segmentations and completeness of recall.

Segmentational Purpose of Cuts	**Definition**
Irrelevance Hypothesis	Viewers ignore cuts as segmenting cues and rely solely on universal breakpoints

Substitution Hypothesis	Viewers consider cuts rather than universal breakpoints as primary segmenting cues
Supplementation Hypothesis	Viewers consider both cuts and breakpoints in parallel
Accentuation Hypothesis	Viewers could attend primarily to breakpoints, but in a more accentuated manner when accompanied by cuts (whereas cuts alone would have no segmentational effect)

Table 1: *This table gives the definitions of fourpotential hypotheses for the Segmentational effect or purpose of film cuts.*

Since edits present transient visual change, even if they go unnoticed due to strong view integration, can cutting at event boundaries subconsciously enhance segmentation at that moment and therefore enhance recall? If this is the case, filmmakers may be able to use this to their advantage, placing important plot points or crucial moments of dialogue at or around these boundaries. Conversely, does purposeful avoidance of cutting at non-event boundaries impair recall? This leads to the question of just how much can the structure and editing of a film control the memorability of the events that take place on screen.

In combination with editing style and what we know about how saccades can be controlled by edits, an argument can be made for film's ability to manipulate one's attention and potentially one's event segmentation to improve subsequent recall of the movie's events.

This thesis will investigate how the editing of a film clip can influence the memorability of movie events. Specifically, we will 1) demonstrate behaviorally the differences in scene recognition and time discrimination accuracy of events improves when clips are edited in accordance with event boundaries, 2) demonstrate the variations of response time in these two tasks across various cut conditions, 3) investigate electrophysiological data to find neurons that indicate these cognitive boundaries and observe how their neural responses change on the basis of film editing.

METHODS

Stimulus Creation

Twenty-five 30-second clips were created that were specifically filmed with the intention of being used for event segmentation in this study. The clips were stripped of sound and consisted of the same five individuals throughout the clips. The events depicted in the clips are day-to-day activities with few to no salient events (Ex:

Folding laundry, writing bills, playing with dog, etc.). Each clip was filmed from three different angles (front view and approximately 45 degrees to either side), using three cameras: Sony Z150, Canon Rebel T3i and an iPhone 8 Camera).

Confirming Event Boundary Locations

Once we created twenty-five 30-second clips each consisting of what appeared to be 3-4 events, we presented them to a group of Amazon MTurk workers (Group 1) to determine the precise locations of the universal event boundaries. The Amazon MTurk workers were recruited under the approval of the Institutional Review Board of Boston Children's Hospital and informed consents were obtained and signed for each subject. Group 1 performed the tasks remotely. These workers (n = 10) were asked to watch the unedited clips, and tasked with indicating each time point where they observed the boundary to a new event (Figure 2). To ensure understanding of the task, we used the following instruction:

> "Please mark off the time point whenever you need to add a new phrase to summarize what just happened in the clip and also type down the phrase you think would best summarize the content you have just seen. Please summarize the clip as concisely as possible. For example, "He turned around, walked over, pushed the door closed, turned around and walked back" can be summarized as "He closed the door". This description is adapted from a previous study: *Darren Newtson, Attribution and the unit of perception of ongoing behavior, Journal of personality and social psychology, 1973."*

As the filmmaker, I first performed the event segmentation task myself, watching an unedited version of each clip, and marking the time points where I observed event boundaries as well, which can be observed as subject 11 in Figure 2.

Figure 2: Individuals Identify Event Boundaries at Similar Times. *Subjects in Group 1 watched all 25 unedited 30 second clips and were asked to indicate when they perceived an "Event Boundary" or transition between one event to another. Subjects 1-10 represent the Event Segmentations of 10 Amazon MTurk workers. Subject 11 is the filmmaker. The blue ticks represent event segmentations that were in response to clip onsets (boundaries between clips) and the black ticks represent event segmentations within a clip.*

Our initial methodology in asking participants to individually segment video clips is adapted from studies by Darren Newtson and his successors, that demonstrated observers can segment streams of action/events into units, separated by boundaries with a high degree of intraindividual reliability and interindividual consistency (Massad, Hubbard, & Newtson, 1979, Newtson 1973, Newtson, Engquist & Bois 1977). Figures 3-4 confirm these claims.

My segmentations were used as a template to confirm the universality of these boundaries. By creating a window of 1.5 seconds (extending ¾ of a second before and after each of the filmmaker's segmentations) we observed an average consistency rate of 89% participants also observing an event boundary in this window (Figure 3). This is 0.5 seconds more strict than the window chosen in Schwann 2000, a similar study regarding edit placement and how they may affect memory formation.

During this task, the clips were observed in succession, therefore also causing what we can refer to as *clip onsets* (a very obvious change caused by the start of an entirely new clip, rather than a within-clip event). These boundaries are obvious and result in near perfect consistency, so they were also used as a measure of how engaged the participants were during viewing. If a participant did not segment during these obvious hard boundaries, we would assume they were not paying attention. Due to the 100% accuracy of those specific boundaries, we removed those segmentations from the consistency rate. Once removed, the resulting consistency of segmentation across the subjects remained high at 86%.

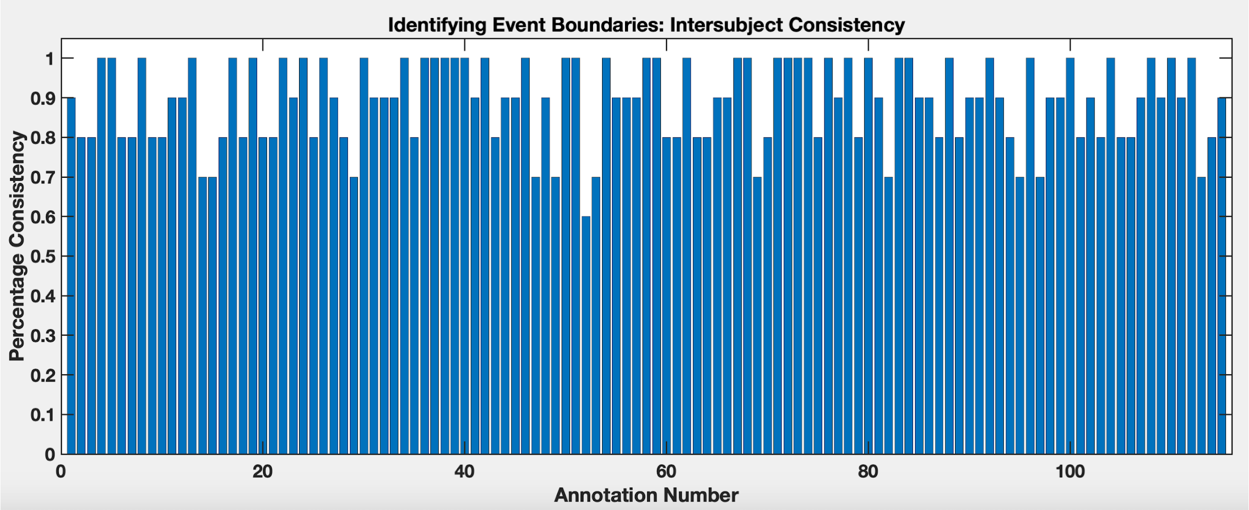

Figure 3: Identifying Event Boundaries: Intersubject Consistency. *Using the filmmaker's segmentations as a guide, we tested consistency of segmentation by extending 0.75 second before and after each of the filmmaker's segmentations to create windows of 1.5 seconds. The Consistency Rate of segmentation within 1.5 second windows across the 11 subjects was 89%. When the 27 clip onset annotations were removed (obvious boundaries between clips with 100% segmentation rate), the Consistency Rate across subjects was 86%.*

We also tested the within-subject consistency of segmentation by repeating three of the clips and comparing where individuals marked event boundaries to their initial segmentations. Figure 4 depicts the average difference in frames between participants' markings the first and second time seeing these three clips. The average within-subject difference across the three clips was around 11-12 frames which, considering the clips were filmed at 24 frames per second, means the average variation in segmentation was about half a second.

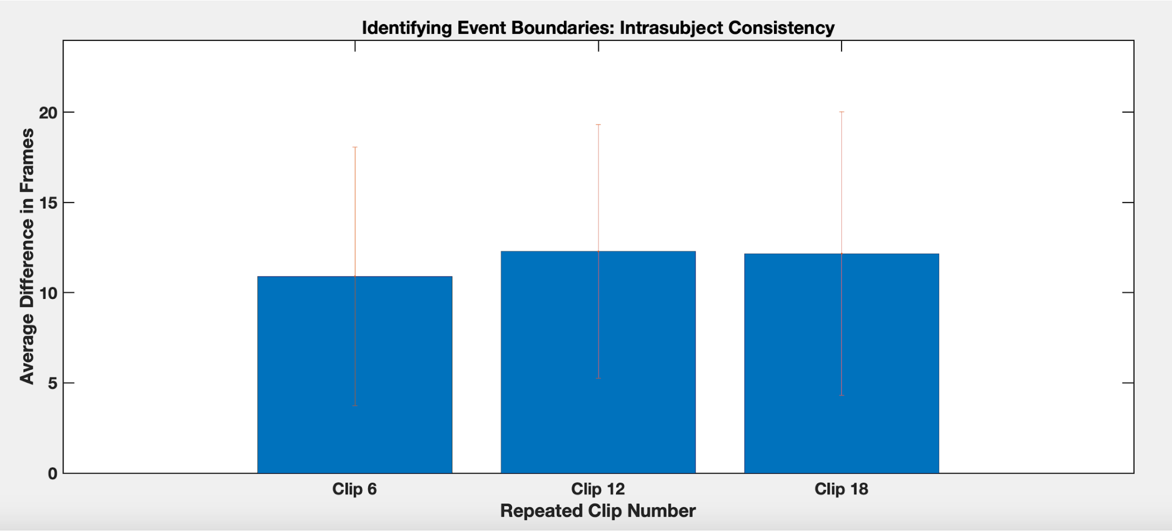

Figure 4: Identifying Event Boundaries: Intrasubject Consistency. *The Consistency of segmentation within subjects was measured by repeating clips 6, 12, and 18. The segmentation times were then compared across the first and second viewing. The average differences in segmentations across the three clips was 11 and 12 frames, which is about half a second, as the film was played at 24 frames per second. Error bars denote standard deviation.*

Creating Three Distinct Cut Conditions

With the event boundaries of the clips identified, we set out to see how the placement of film edits with regard to these boundaries could impact episodic memory formation. We created three different conditions for the 25 clips (Figure 5). The Good Cut condition (GC) refers to when edits (change in camera angle) were placed where event boundaries were observed (The filmmaker's segmentation pattern was used as the guide, because of results presented in Figure 3 showing an 86% consistency rate in those time windows) . The Bad Cut condition (BC) will refer to when edits occurred mid-event, purposefully avoiding event boundaries. These edits cut across the motion of the event. Lastly, the No Cut condition (NC) was created as a control group, to compare our results to a baseline reaction to events without any edits.

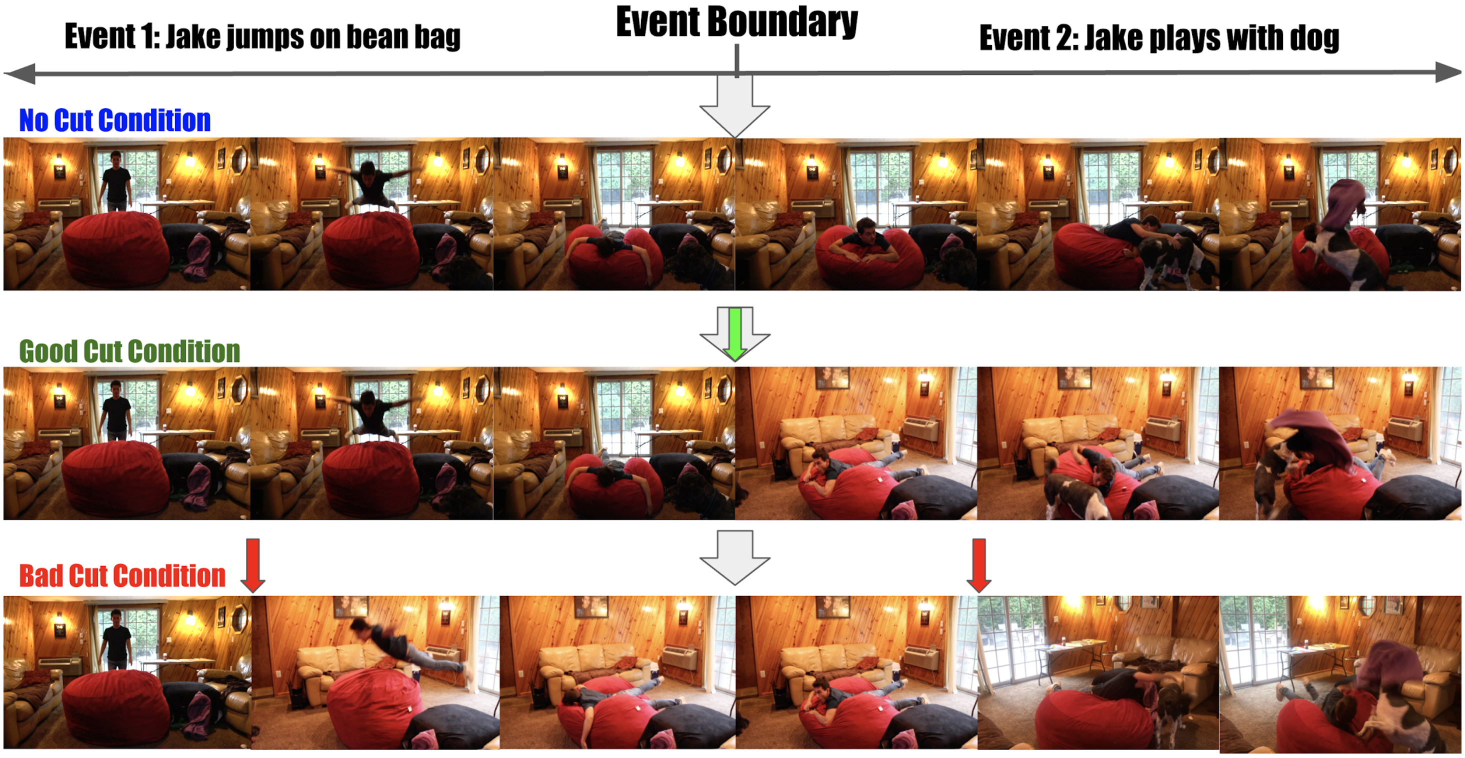

Figure 5: Cut Conditions Diagram. *Each of the three conditions depicted the same events but the location of the cut (when the camera angle changes) differed. Note the gray arrow indicates the location of the event boundary in the clip. The No Cut condition served as a control case, consisting as one continuous angle. The Good Cut condition contained a cut at the point of the event boundary (indicated by the green arrow). The Bad Cut condition cut across motion, during points in the events that were not the event boundary (indicated by the red arrows).*

Experimental Tasks

The rest of the research was conducted across an independent group of 18 subjects. The first 15 of these subjects were another independent group of Amazon MTurk workers (Group 2), used only to obtain behavioral data. The last three subjects were drug-resistant epilepsy patients who were implanted with hybrid depth electrodes from which we obtained single-unit neuronal data as well as behavioral data.

The encoding stage involved watching all 25 clips, which consisted of a mixture of GC, BC, and NC versions. To ensure the subjects remained engaged during this encoding process, we asked a simple true/false question intermittently, about every 3-5 clips (Example: Was anyone in the clip wearing glasses? Did anyone in the clip have long hair? *Overall Accuracy* = 76.75% ± 4.66%).

After viewing the clips, the subjects were prompted with two memory tasks. The first was a Scene Recognition task where they were shown a movie frame and asked if it was part of the stimuli they had encoded. They were first shown a fixation cross for approximately 0.5 seconds, then shown the image, and shortly after, two words, “Old” and “New” appeared on the screen from which they would choose what they believed to be the correct answer. We incorporated a range of difficulties in the various foil frames we presented across 105 trials, ranging from easy (images from other projects that very clearly had not been a part of the stimulus set) to very difficult (a frame with the same character and location as seen in the clips, depicting an event that didn’t actually happen during encoding).

The second task was a Time Discrimination task where the subjects were first shown a fixation cross (again for about 0.5 seconds), and then presented with two frames from a clip they had observed during encoding. These frames were always separated by one or more event boundaries, regardless of what cut condition (GC, BC, or NC) the subject observed during encoding for that clip. They were asked to identify which of the frames (left or right) was presented first, across 166 trials.

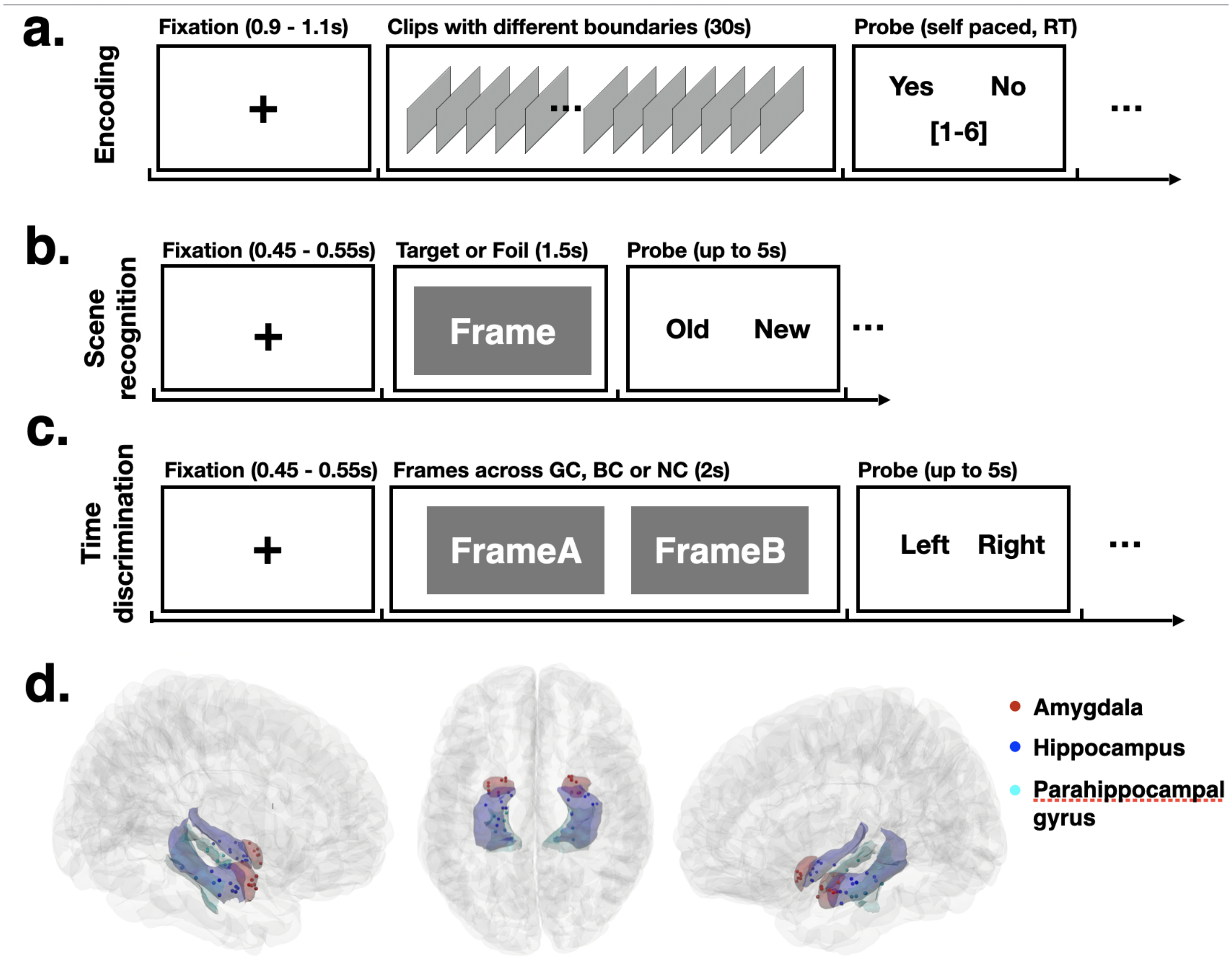

Figure 6: Experimental Design Schematic. *a.) During the encoding stage, a fixation cross was presented before each clip onset, followed by the 30 second clip. A total of 25 clips were watched containing 8 GC, 8 BC, and 9 NC condition clips. Intermittent questions were asked every 3-5 clips to ensure attention. (Example: Was anyone in the clip wearing sunglasses?; Overall Accuracy = 76.75% ± 4.66%). b.) The Scene Recognition task involved a fixation cross and the presentation of either a target or foil frame from which the subject would choose "OLD" or "NEW". c.) The Time Discrimination task involved a fixation cross as well, and the presentation of a pair of two frames across one or more boundaries. The subjects were tasked with determining which frame came first chronologically, frame LEFT or RIGHT. d.) This diagram shows the areas of the brain we are considering when we reference the Medial Temporal Lobe where the electrodes were placed: the amygdala, hippocampus, and parahippocampal gyrus.*

Electrophysiology

The neural patients were drug-resistant epilepsy patients who volunteered for this study and gave their informed consents. The institutional review boards of Toronto Western Hospital and Cedars Sinai Medical Center approved all protocols. The locations of the implanted electrodes were based on clinical needs alone, and the tasks were completed while the patients stayed in the hospital as they were being monitored for seizures.

We recorded bilaterally in patients with drug-resistant epilepsy using hybrid depth electrodes (Ad-Tech Company, Oak Creek, Wisconsin, USA). They contained eight 40-µm diameter microwires at the tip of each electrode shank. Each microwire recorded at 32kHz using the ATLAS system (Neuralynx Inc., Bozeman, Montana, USA) through broadband signals (0.1-9000 Hz filtered). The neuronal signals were filtered (zero-phase lag) offline in the 300 to 3000 Hz band. We detected/sorted spikes using a semiautomated template matching algorithm called Osort (Friend & Rutishauser, 2014; Rutishauser 2006). The resulting data was analyzed and primarily deemed significant using t-tests with a threshold of $p < 0.01$.

RESULTS

After analysis of the scene recognition data across all 18 subjects, there was no significant difference observed in scene recognition accuracy across the three cut conditions (*GC Accuracy* = 75.2% ± 8.55%, *BC Accuracy* = 75.2% ± 8.49%, *NC Accuracy* = 74.7% ± 7.76%, *p* = 0.9764).

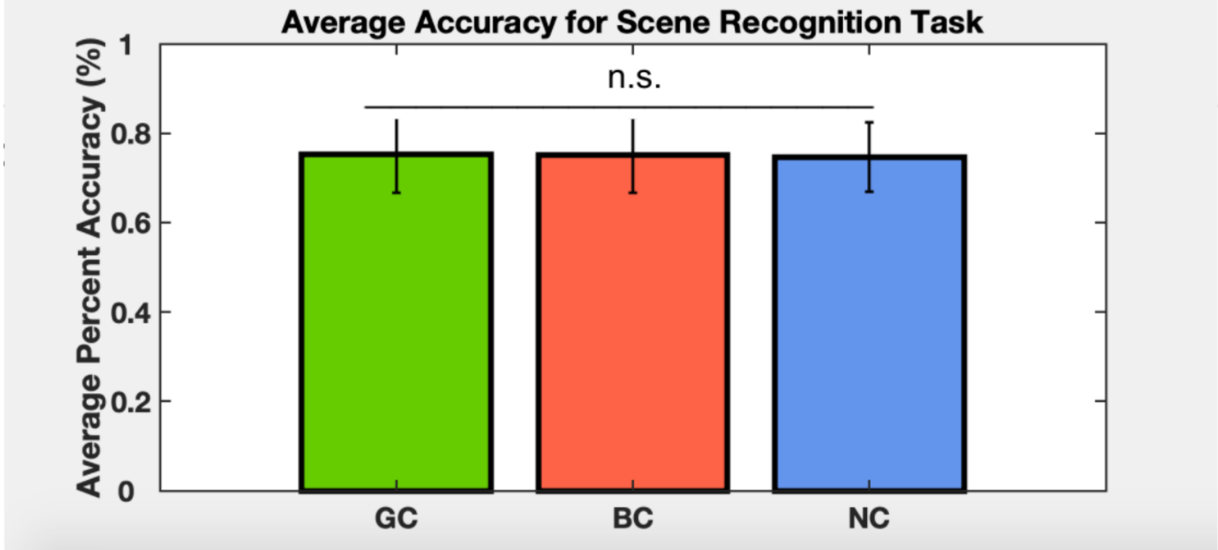

Figure 7: Average Accuracy for Scene Recognition Task. *The average percentage accuracy for the Scene Recognition Task across all patients (n=18) was not significantly different (t-test, p=0.9764) when comparing across the three cut conditions (mean± std; GC Accuracy = 75.2% ± 8.55%, BC Accuracy = 75.2% ± 8.49%, NC Accuracy = 74.7% ± 7.76%).*

Similarly, there was no significant difference in response time across the three conditions (*GC Response Time* = 1.56 ± 0.21 seconds, *BC Response Time* = 1.53 ± 0.21 seconds, *NC Response Time* = 1.55 ± 0.19 seconds, $p = 0.9761$).

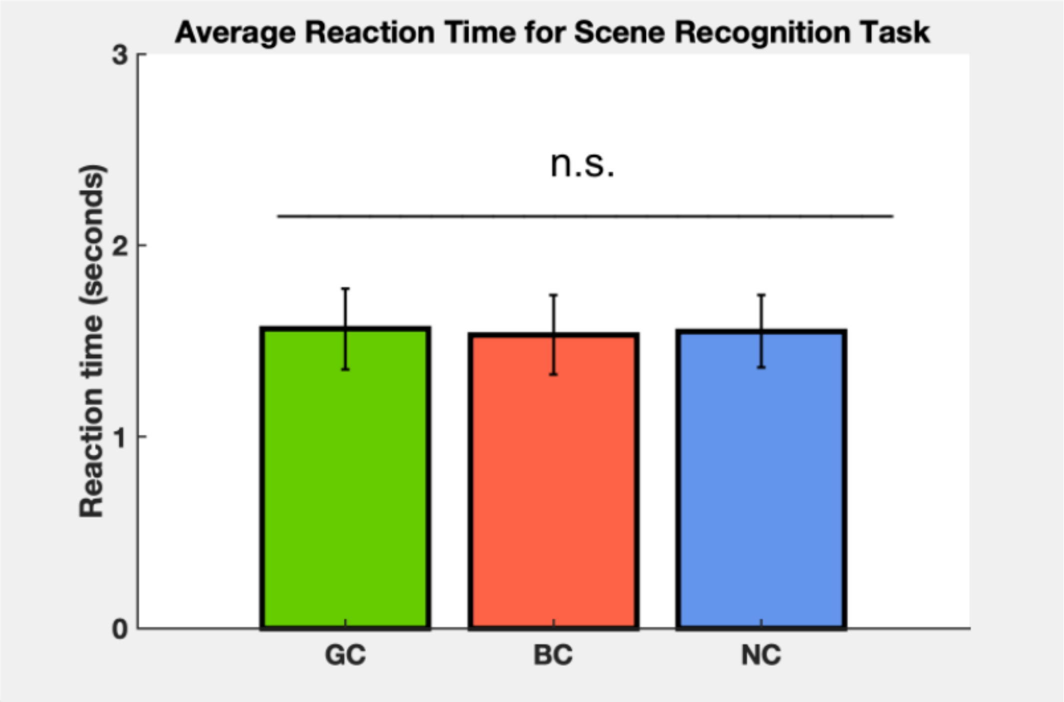

Figure 8: Average Reaction Time for Scene Recognition Task. *The average response time for the Scene Recognition Task across the patients (n=18) was not significantly different (t-test, p=0.9761) when comparing across the three cut conditions (mean± std; GC Response Time = 1.56 ± 0.21 seconds, BC Response Time = 1.53 ± 0.21 seconds, NC Response Time = 1.55 ± 0.19 seconds).*

We observed that the average time discrimination accuracy for GC clips was significantly greater than BC and NC clips (*GC accuracy* = 71.5% ± 2.0%, *BC accuracy* = 65.6% ± 2.8%, *NC Accuracy* = 66.5% ± 2.9%, $p = 9.41\times10^{-9}$).

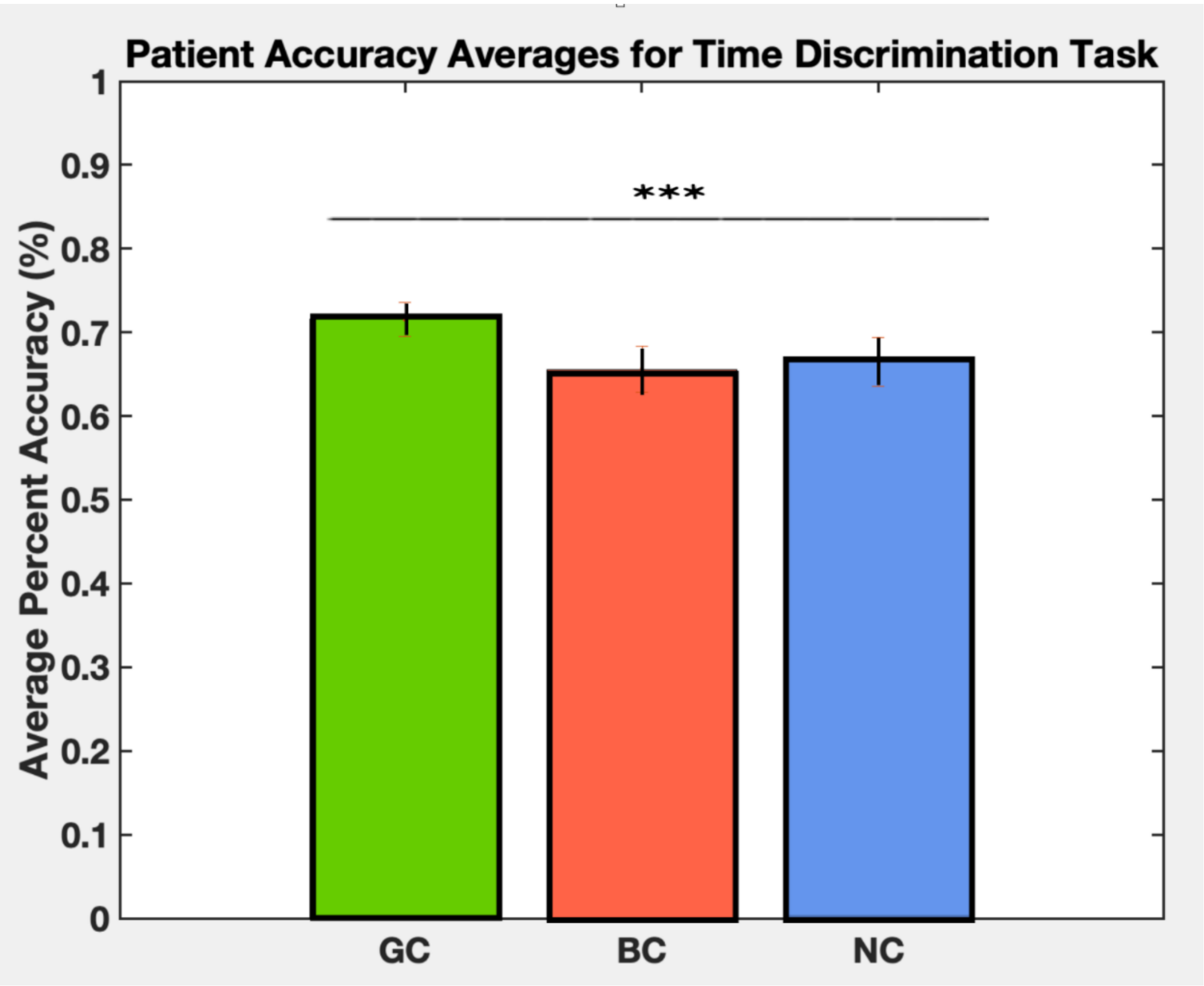

Figure 9: Patient Accuracy Averages for Time Discrimination Task. *The average percent accuracy for the Time Discrimination Task across the patients (n=18) had a ~5% significant difference for the GC condition (t-test,* $p = 9.41\text{x}10^{-9}$*) when comparing across the three cut conditions (mean± std; GC accuracy = 71.5% ± 2.0%, BC accuracy = 65.6% ± 2.8%, NC Accuracy = 66.5% ± 2.9%).*

Interestingly, the average response time for this Time Discrimination task was significantly longer for the BC clips than the GC and NC conditions (*GC Response time* = 1.15 ± 0.23, *BC response time* = 1.54 ± 0.31 seconds, *NC Response Time* = 1.19 ± 0.18 seconds, $p = 2.18\text{x}10^{-5}$).

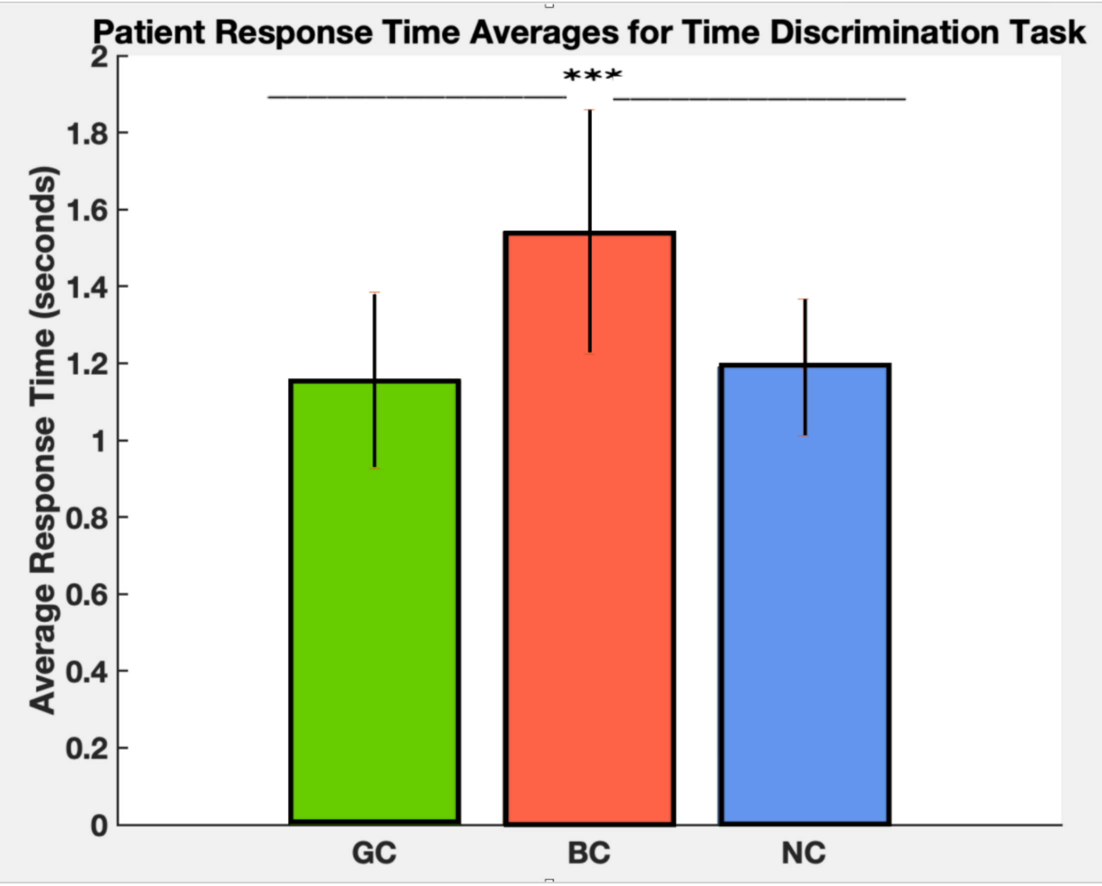

Figure 10: Patient Response Time Averages. T*he average response time across the patients (n=18) for the Time Discrimination Task had a significant difference for the BC condition (t-test, p =* $2.18\mathrm{x}10^{-5}$*) when comparing across the three cut conditions (mean± std; GC Response time = 1.15 ± 0.23, BC response time = 1.54 ± 0.31 seconds, NC Response Time = 1.19 ± 0.18 seconds).*

We next analyzed the single-unit neural responses recorded during encoding from our drug-resistant epilepsy patients. Figure 11 includes a diagram of the electrode localization for patients 148 and 149.

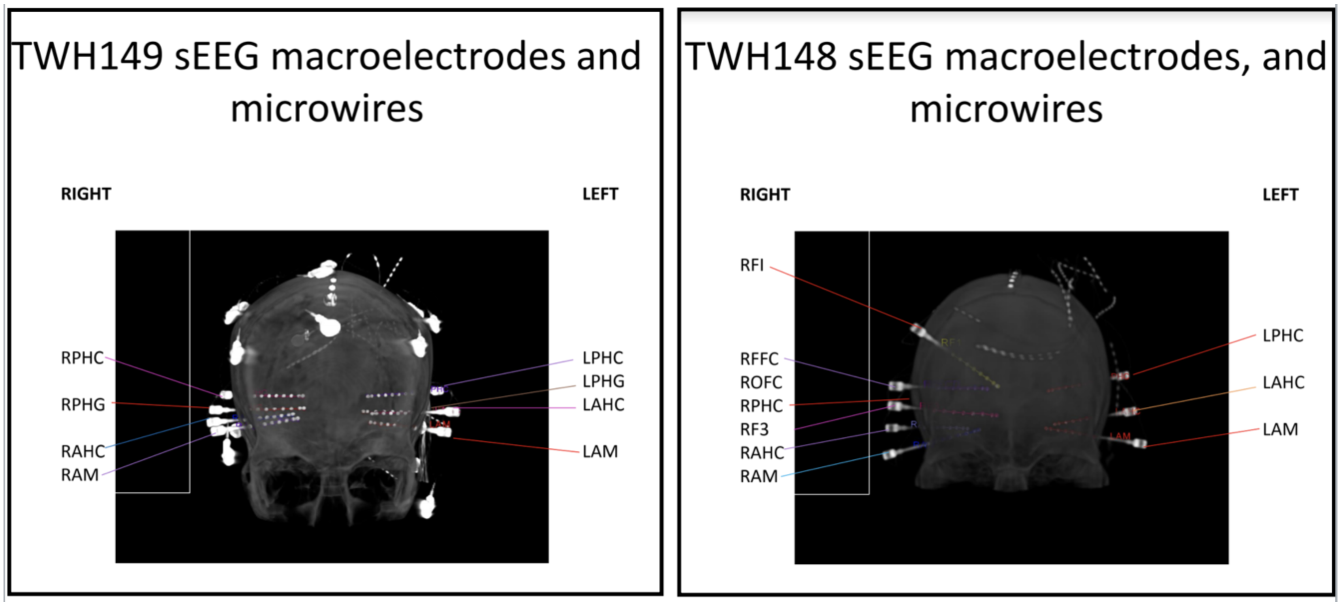

Figure 11: Electrode Localization Diagram. *Across our subjects, we identified 147 neurons from which we could analyze variation in firing patterns, recorded from the following areas: Subject TWH149 (left) had electrode shanks implanted in the Right Posterior Hippocampus (RPHC), Right Parahippocampal Gyrus (RPHG), Right Anterior Hippocampus (RAHC), Right amygdala (RAM), Left Posterior Hippocampus (LPHC), Left Parahippocampal Gyrus (LPHG), Left Anterior Hippocampus (LAHC), and Left Amygdala (LAM). Subject 148 (right) had electrode shanks implanted in similar locations with the addition of the Right Frontal Insula (RFI), Right Orbitofrontal Cortex (ROFC), Right Frontal Cortex (RF), Right Fusiform Face Complex (RFFC).*

To test whether edit placement within clips could affect firing rate, we created raster plots and peristimulus time histogram of firing during encoding. We aligned the plot and histogram to cuts across the Good Cut, Bad Cut and No Cut conditions. We evaluated neurons' responses to the three cut conditions by analyzing their spike rates 0.5 seconds after cuts relative to baseline (0.5 seconds before cuts). We set a threshold difference of $p < 0.01$ to be considered having conditional significance. In the case of the NC conditions, we aligned responses to the 0.5 seconds before and after the event boundary location (the same location of GC edits, simply without a cut). Figure 12 shows an example of two neurons we found that exhibited a significant increase in firing rates after both good cuts (green) and bad cuts (red). No such change was observed in the NC conditions. We referred to this type of neuron as a *General Cut Responsive Cell*.

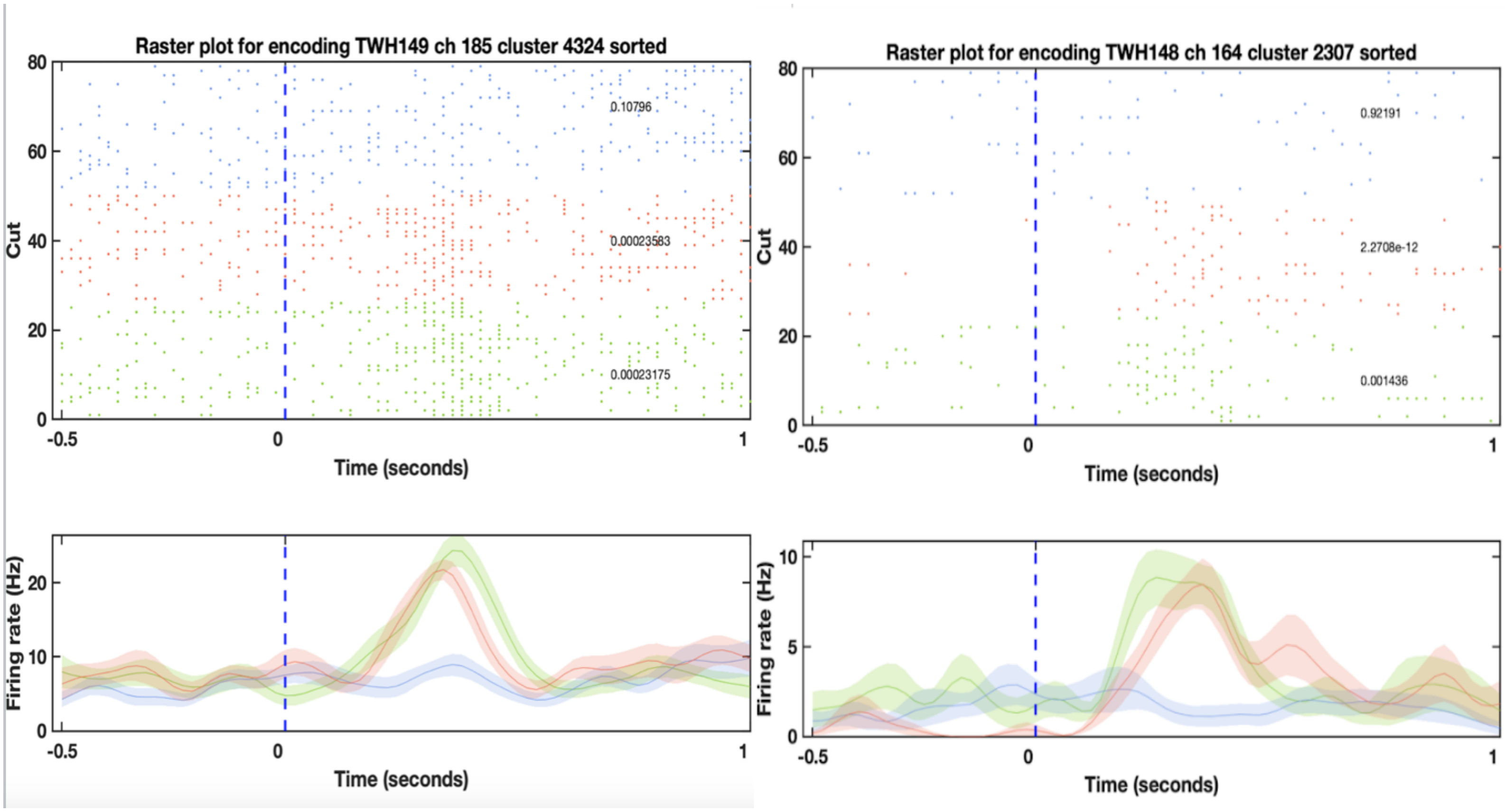

Figure 12: Example of General Cut Responsive Cell. *Above depicts a neuron from Patient TWH149 (left) and Patient TWH148 (right). These two neurons were identified as a General Cut Responsive Cell. The dotted blue line is aligned to cuts in each of the clips for GC (green), BC (red), and NC (blue). Note the NC raster plot acts as a control to GC, aligned to event boundaries despite no cut being present in that condition. A T-test comparing the sum of spikes 0.5 seconds before and after the cut locations produced p-values that we used to evaluate the response of the neuron. This neuron responded significantly to GCs and BCs, but not to NCs. The peristimulus time histogram below plots the firing rates before and after cuts. A large increase in firing rate immediately after cuts can be observed for the GC and BC condition in both of these neurons.*

Was this neuron simply responding to any transient stimulus change? The above figure shows a significant increase in firing for both conditions with cuts but not in the condition with no cuts. We looked at the same two neurons' firing rates aligned to clip onset in order to see if this neuron was also responding significantly to this hard boundary between clips rather than just the within-narrative event boundaries. Figure 12 shows these neurons did not respond significantly to clip onsets.

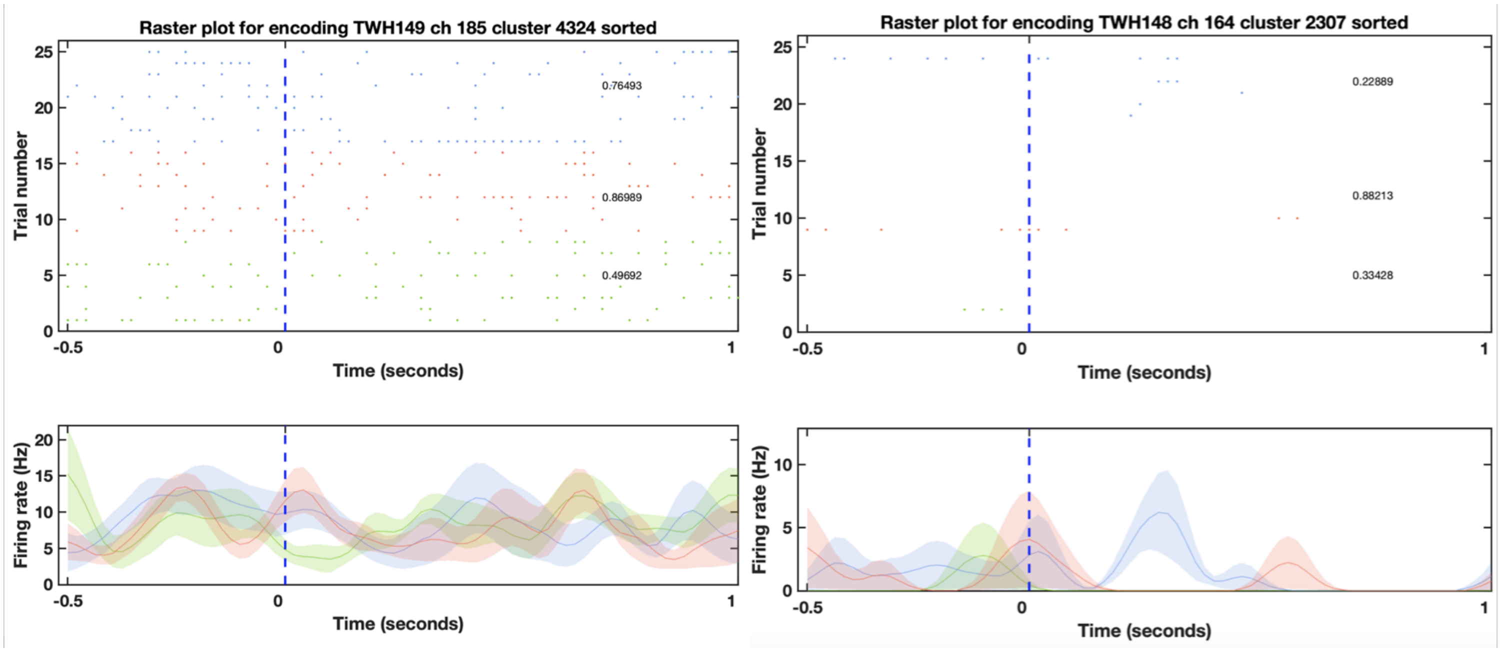

Figure 13: General Cut Responsive Cells do not respond to Clip Onsets. *The two General Cut Responsive Cells did not respond significantly to clip onsets, suggesting that its response is not generalized to any stimulus change, but specific to within-clip stimulus change.*

We also identified 6 neurons which we dubbed *Good Cut Responsive Cells.* In these neurons, there was a transient increase in firing only after cuts that coincided with event boundaries (GC). These neurons did not respond to edits placed during events (BC) or to event boundaries when there was no edit (NC). However, when pairing an edit with an event boundary, these 6 neurons were triggered significantly (Figure 14).

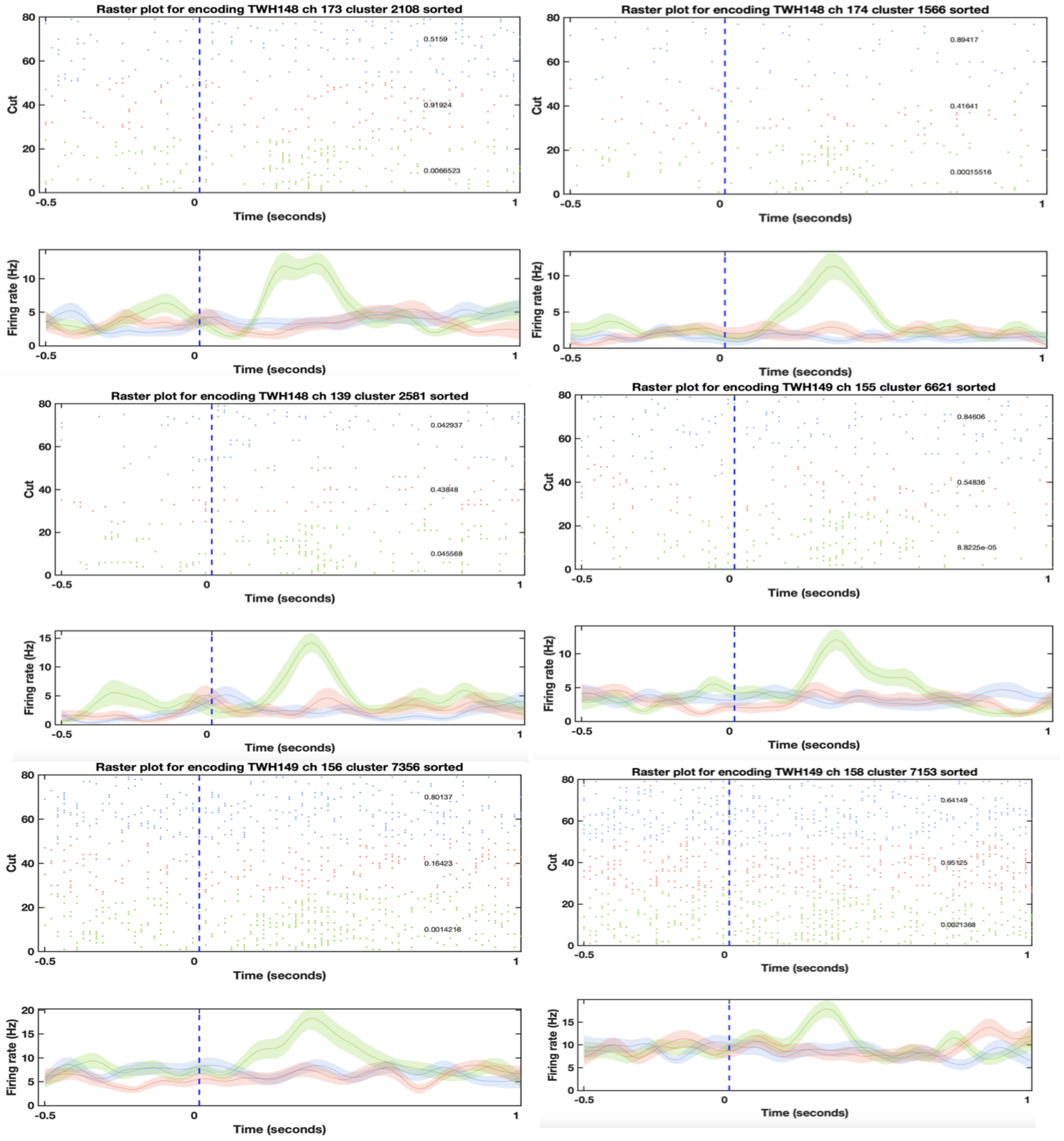

Figure 14: Good Cut Responsive Neurons. *Shown above are three neurons from subject TWH148 and three from TWH149 that exhibit a significant increase in firing rate (t-test, p<0.01) in response to edits paired with event boundaries (Good Cuts).*

Interestingly, we also analyzed the neurons by splitting up the data by cut chronology rather than averaged across the cuts in the clip. Although more data is needed to make more conclusive statements, we found the *Good Cut Responsive Neurons* each exhibited a preference for specific cut (Figure 15).

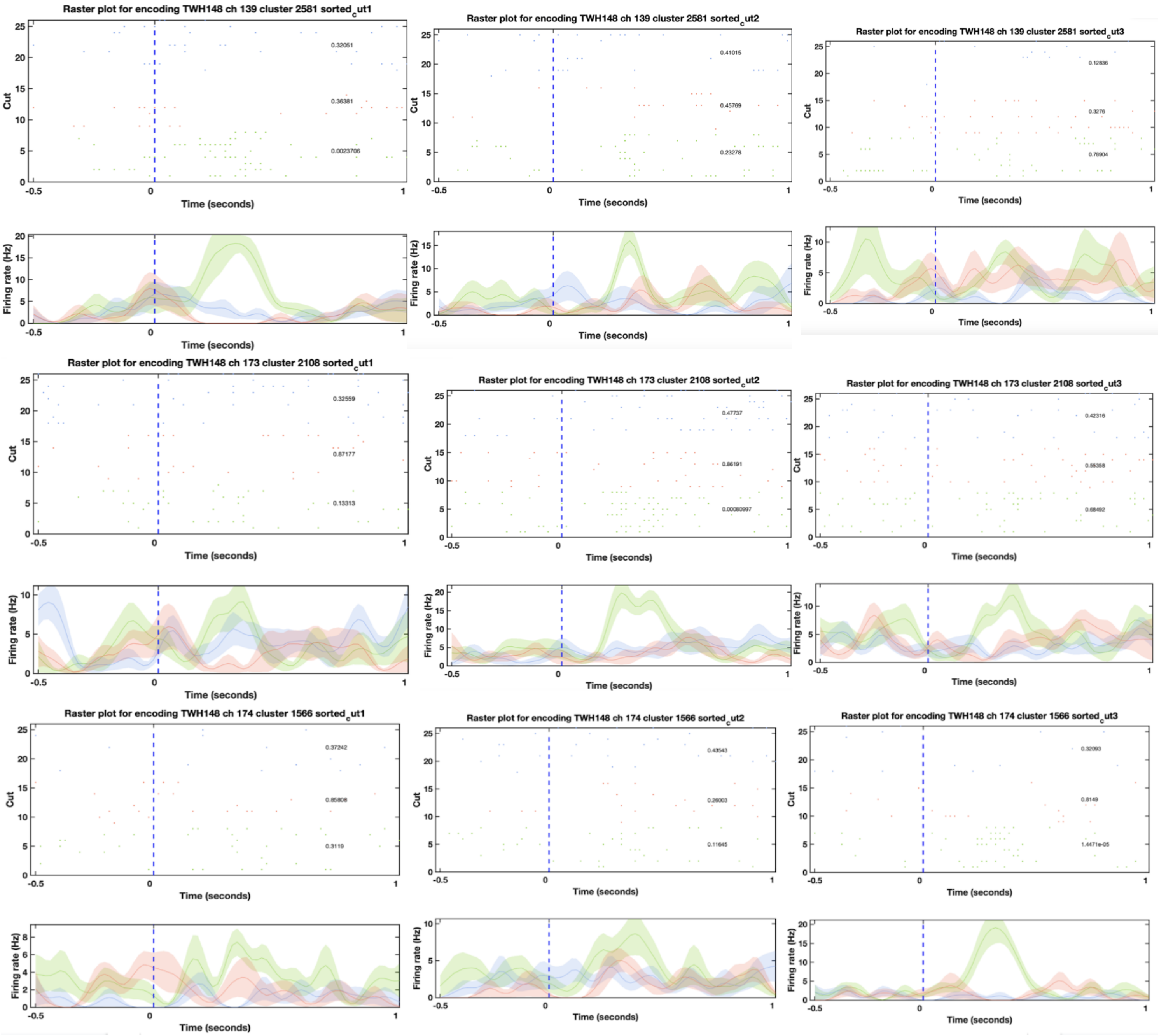

Figure 15: Good Cut Responsive Neurons with Strong Cut Preferences. *Above depicts three Good Cut Responsive Neurons from patient TWH148 that each respond most significantly to one of the three cuts in a given clip for the Good Cut condition. Row 1, from left to right, shows the spike in firing for this neuron in cut 1 ($p < 0.01$), and less significant spikes for cuts 2 and 3. Row 2 shows a different neuron from TWH148 that exhibits a significant increase in firing for Cut 2 ($p<0.01$). Finally, row 3 represents a third neuron that indicates a preference ($p<0.01$) for cut 3.*

The remaining 3 of 6 Good Cut Responsive neurons not included in Figure 15 were observed in patient TWH149, two of which exhibited a preference for Cut 3 and one for Cut 2 ($p<0.01$).

DISCUSSION

We observed that the location of edits had no effect on scene recognition accuracy or response time, but did improve Time Discimination accuracy in the GC condition and lengthened response time for the BC condition. We also discovered two General Cut Responsive Neurons that responded significantly to both GC and BC conditions, as well as six Good Cut Responsive Neurons, which each responded particularly strongly to a specific cut. These results seem to support our hypothesis that placing edits at event boundaries facilitates memory formation, at least in the recall of temporal structure, and that the coincidence of an edit with event boundaries incites a specific reaction from Good Cut Responsive Neurons. Our results also invite further questions having to do with if these cut-selective Good Cut Responsive Neurons are being used to store the chronology of events in the brain.

This thesis builds off of work done by many researchers in the pursuit of furthering the understanding of the brain's mechanism for event perception, subsequent memory formation, and of course, how movies are perceived, beginning with our stimulus creation. We created a more controlled stimulus set than previous studies with similar hypotheses like Zheng, 2019, which used stimuli from professional Hollywood movies. By creating stimuli specific for this study with a set length, limited characters, no sound, and being directed with event segmentation in mind, we created a more controlled environment to study this phenomenon. Our clips were purposefully day-to-day tasks with very few to no salient events versus a Hollywood movie whose purpose is to entertain. A potential limitation however was the use of three different types of cameras (Sony Z150, Canon T3i, and iPhone camera), which had inherent differences in quality such as light exposure and lens length. These differences may have caused an additional transient change in exposure or composition that was unintentional. This may have given rise to more noticeable boundaries at cuts than the same scenario where both clips were filmed by an identical camera and lens.

Our results (Figures 7 and 8) suggest that no matter where the edits were placed in the clips, subjects recognize scenes with just about equal accuracy and swiftness regardless of cuts. Schwann (2000) asked a similar question, pairing cuts with event boundaries and performing a scene recognition task in hopes of observing greater accuracy and also did not observe any effects. These hypotheses came from previous literature which recognizes that event boundaries tend to be remembered disproportionately well (Perrijohn et. al 2016, Radvansky et. al 1998, Huff et. al, 2014). Another study, Newston and Enquist (1976) suggested that event boundaries may be the building blocks for episodic memory, and therefore greater resonance of segmentation (coinciding with a film cut) may increase recall for that event. Our results as well as the results from Schwann (2000) do not support that claim.

However, we observed a statistically significant 5% increase in accuracy for the GC conditions during our time discrimination task (Figure 9). Schwann (2000) made a similar claim with conditions where cuts coincided with event boundaries, through a written recall task. Participants wrote down a detailed ordered protocol of the events they observed as part of the memory task. This study observed more detailed protocols in response to watching clips with edits that coincided with event boundaries. Their results in combination with our increase in time discrimination accuracy for the GC condition seems to suggest that when edits coincide with event boundaries there is an enhanced ability in recalling the temporal structure of events.

A similarly interesting result was the longer response time within BC conditions during the time discrimination task. This seems to suggest that cutting across motion, as the BC edits were placed in the middle of events, makes retrieval of the event chronology take slightly longer. This result is particularly interesting when considering canonical film editing. Although in this thesis, we refer to them as “bad cuts”, filmmakers use these cuts very often, dubbing them “match-action” cuts. They are particularly beloved by filmmakers because they appear so smooth and work to keep the spectator engaged in the narrative content instead of the low-level visual changes (Zacks, 2015). As many as a third of these edits go unnoticed when participants are asked to attend to them (Smith, 2012). The GC condition in our study very rigidly places edits at the boundary of events, and this supplement of pixel-change at the boundary of an event may facilitate segmentation and subsequent chronological or detailed recall, but results in a more abrupt, noticeable change.

Additionally, we found two General Cut Responsive Cells that responded significantly to both GC and BC conditions, and did not have a significant response to NC conditions or to clip onsets. This suggested that these neurons had a conditional significance to cuts across a continuous narrative. As discussed above, transitioning from continuous experience to discrete events is said to be highly reliant on cognitive boundaries. They help to extract structural information of experience. These preliminary results can be discussed in conversation with rodent studies which have located boundary cells in the hippocampal subiculum that fire at edges (Lever 2009; Stewart, 2014; Bird, 2010). Other studies have found that when rodents cross to new sections of space, hippocampal cell firing rates spike. (Spiers, 2015). These spatial boundaries mark the structure of continuous spatial experience, similar to how firing patterns may be reflected at event boundaries. Whittington (2020) offers a link between spatial and relationship memory tasks, claiming that a unified network may be responsible for obtaining the structure knowledge for spatial and non-spatial scenarios. With this in mind, cells like these General Cut Responsive Cells may be analogous to the types of boundary cells discovered in spatial rodent studies.

The 6 Good Cut Responsive Cells we discovered seem to suggest that pairing an edit with an event boundary is enough to cause these neurons to react significantly, whereas just the event boundary (NC condition) does not suffice. Interestingly, we found that of the six Good Cut Responsive Cells, they each respond most strongly ($p<0.01$) to a particular cut. This may be a neural representation of storing the chronology of events. Sun and Yang (2020) showed that rodents have different hippocampal neurons that respond to event laps. This spatial-lap encoding may be the spatial equivalent to the Good Cut Responsive Neurons that seem to indicate preference for one of the three cuts. With further research, it may be possible to demonstrate that these Good Cut Responsive Neurons also provide a temporal structure of events based on firing with specificity to the chronology of events.

These neurons are unique in comparison to Zheng (2019) which found neurons called *Boundary Cells*. These cells responded to within clip cuts as well as clip onsets (in their study clip onsets are referred to as *hard boundaries)*. This study also found what they referred to as *Event Cells* which only responded to hard boundaries. These neurons were found in the parahippocampal gyrus, the amygdala, and the hippocampus. Our Good Cut Responsive Cells exhibited additional specificity, as no response was exhibited in response to clip onsets (hard boundaries) or to just any within clip cut (BC). The cut had to be within clip and placed over an event boundary (GC) to fire.

Zheng (2019) further analyzed the neural responses during encoding and linked the spiking to the behavioral results. They theorized that the response of the boundary and event cells during encoding may predict the subjects' subsequence performance in their two tasks (similar scene recognition and time discrimination tasks). They measured whether the strength of responses of the cells they found varied as a function of whether the familiarity or temporal order of a stimulus was remembered or forgotten later. They found that the response to boundaries that preceded frames that were remembered was stronger for boundary cells in their scene recognition task ($n = 42$). This was not the case for boundary cells in the time discimrination task. This was also not the case for event cells in the scene recognition nor the time discrimination task ($n = 36$). However, it is important to note that the Zheng study only had one cut in each clip, dividing just two events, so our study, consisting of a data set with 3-4 events in each clip, would have a stronger effect for structural information. From this thesis, we hope that future analysis of the collected data will investigate how our General Cut Responsive Cells and Good Cut Responsive Cells fired in relation to the correct and incorrect trials of our memory tasks. The hypothesis would be that Good Cut Responsive Cells had a stronger firing rate for the corerct trials of the time discrimination tasks, providing a link between increased neuronal firing in the MTL and enhanced recall of temporal strucutre. Judging from our behavioral accuracy results, we would not hypothesize a similar connection in the scene recognition trials, as the behavioral data showed no significant differences, but the Zheng study suggests that firing rates may vary enough to provide for a predictive relationship despite there not being a difference in accuracy across cut conditions.

It is important to note, also, that this is a pilot study and further analysis as well as an increase in participants would greatly help solidify the legitimacy of our results. Recording from only 18 subjects makes all of our results less reliable than if we were to recruit a much larger set of participants in the hundreds. Additionally, analysis could be improved by having more neural subjects, so that the behavioral analyses could be run separately on MTurk Workers and neural subjects to ensure that there is not a difference between subject groups. On a similar note, although single-unit data collection often involves low numbers of subjects, recording from a larger number of neurons would give us more confidence in the types of neurons we observed with a greater possibility of analysis that the results we are observing are robust and replicable.

This thesis merges two very different fields, neuroscience and filmmaking, and seeks to use the strengths of each to benefit the other: Neuroscientific research to make more engaging cinema, and films as a means to understand the brain. The making of art and the practice of science together is untraditional and often avoided. However, this thesis works to exemplify an interest both of these fields have in common: Event Perception. Filmmakers intend to transport their viewers into their stories as effectively as possible, and therefore work hard to achieve perceptual continuity and conceptual integration, leaning on match-action shots to go unnoticed and satisfying the visual processing stream in ways that will help bring the narrative to life (Hasson, 2008). So, pairing edits solely on top of event boundaries, as in the GC conditions, as a hard and fast rule of editing would hinder the storytelling process. Simultaneously, filmmakers are always striving for better ways to direct their viewers' attention or to provide memorable clues in support of a narrative. Additionally, in some aspects of entertainment, success relies entirely on how memorable and engaging the content is. With this in mind, an understanding of the benefits of a "good cut" may be helpful information for a filmmaker who is looking for an additional way to accentuate an event or detail. Filmmakers are constantly making and breaking rules in order to play with the experience of watching a movie, and unlike scientific research, is often thought of as less formulaic. Many artists believe that thinking too critically about the artistic process hinders expression. However, the results of a study like this can only be helpful information to a storyteller about the cognitive experiences of their viewers based on edit placement, and can help inform their artistic decision making.

References

Abelson RP. 1981. Psychological status of the script concept. *Am. Psychol.* 36:715–29

Baldassano, C. et al. (2017) Discovering event structure in continuous narrative perception and memory. *Neuron* 95, 709-721.

Ben-Yakov, A. and Dudai, Y. (2011) Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *J. Neurosci.* 31, 9032-9042.

Ben-Yakov, A. et al. (2013) Hippocampal immediate post stimulus activity in the encoding of consecutive naturalistic episodes. *J. Exp. Psychol. Gen.* 142, 1255-1263.

Bezdek, M. A., Foy, J. E., & Gerrig, R. J. (2013, September 16). "Run for It!": Viewers' Participatory Responses to Film Narratives. Psychology of Aesthetics, Creativity, and the Arts. Advance online publication. doi: 10.1037/a0034083

Bird, C.M. *et al.* (2010) Establishing the boundaries: the hippocampal contribution to imagining scenes. *J. Neurosci.* 30, 11688-11695.

Brunec, I.K. et. al (2018) Boundaries Shape Cognitive Representations of Spaces and Events *Trends in Cognitive Sciences*, 22, 7, 637-650.

Buzsaki G, Moser EI (2013) Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat Neurosci* 16: 130-138.

Carroll, J.M., & Bever, T. G. (1976). Segmentation in cinema per- ception. *Science,* 191, 1053-1055.

Chen, J. et al. (2017) Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* 20, 115-125.

Cutting, James & Candan Simsek, Ayse. (2013). Movies, Evolution, and Mind: From Fragmentation to Continuity. The Evolutionary Review. 4. 25-35.

Daw, N.D. et al. (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neuroscience* 8, 1704-1711.

Dragoi G, Buzsaki G (2006) Temporal encoding of place sequences by hippocampal cell assemblies. Neuron 50:145-157.

Eichenbaum H (2004) Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* 44: 109-120.

Eisenberg ML, Zacks JM, Flores S. 2018. Dynamic prediction during perception of everyday events. *Cogn. Res. Princ. Implic.* 3:53

Flanagan JR, Johansson RS. 2003. Action plans used in action observation. *Nature* 424:769–71

Foster DJ, Wilson MA (2007) Hippocampal Theta Sequences. Hippocampus 17: 1093-1099.

Gupta AS, van der Meer MAA, Touretzky DS, Redish AD (2012) Segmentation of spatial experience by hippocampal theta sequences. Nat Neurosci 15: 1032 - 1039.

Hafting, T. et al. (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436, 801-806.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, *28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

Hayhoe M, Ballard D. 2005. Eye movements in natural behavior. *Trends Cogn. Sci.* 9:188–94

Hemeren PE, Thill S. 2011. Deriving motor primitives through action segmentation. Front. Psychol. 1:243

Huff M, Meitz TGK, Papenmeier F. 2014. Changes in situation models modulate processes of event perception in audiovisual narratives. *J. Exp. Psychol. Learn. Mem. Cogn.* 40:1377–88

I. Friend, U. Rutishauser, M. Cerf, G. Kreiman, *Single neuron studies of the human brain : probing cognition*. (The MIT Press, Cambridge, Massachusetts, 2014), pp. Viii, 365 pages.

Janzen, G. and van Turennout, M. (2004) Selective neural representation of objects relevant for navigation. *Nat. Neuroscience* 7, 673-677.

Janzen G. et al. (2007) Neural representation of navigational relevance is rapidly induced and long lasting. *Cereb. Cortex* 17, 975-981.

Kraft, R. N. (1986). The role ofcutting in the evaluation and retention offilm. *Journal ofExperimental Psychology: Learning, Memory.* & *Cognition,* 12,155-163.

Lever, C. et al. (2009) Boundary vector cells in the subiculum of the hippocampal formation. *J. Neurosci.* 29, 9771-9777.

Levin, D. T., & Baker, L. J. (2017). Bridging views in cinema: a review of the art and science of viewintegration. Wiley Interdisciplinary Reviews: Cognitive Science , n/a -n/a. https://doi.org/10.1002/wcs.1436

Loschky LC, Larson AM, Magliano JP, Smith TJ. 2015. What would Jaws do? The tyranny of film and the relationship between gaze and higher-level narrative film comprehension. *PLOS ONE* 10:e0142474

Loucks J, Pechey M. 2016. Human action perception is consistent, flexible, and orientation dependent. *Perception* 45:1222–39

Magliano, J.P. and Zacks, J.M. (2011), The Impact of Continuity Editing in Narrative Film on Event Segmentation. Cognitive Science, 35: 1489-1517. doi:10.1111/j.1551-6709.2011.01202.x

Magliano JP, Radvansky GA, Forsythe JC, Copeland DE. 2014. Event segmentation during first-person continuous events. *J. Cogn. Psychol.* 26:649–61

Massad, C. M., Hubbard, M., & Newtson, D. (1979). Selective per- ception of events. *Journal of Experimental Social Psychology,* IS, 513-532.

Munsterberg, H. (1916). *The Photoplay: A psychological study.* New York, NY: D. Appleton.

Murch, Walter. *In The Blink of An Eye: A Perspectiveon Filmmaking.* 2nd Edition. Silman-James Press 1992.

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology,* 12 436-450.

Newtson D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality* & *Social Psychology,* 28, 28-38.

Papafragou A. 2010. Source–goal asymmetries in motion representation: implications for language production and comprehension. *Cogn. Sci.* 34:1064–92

Pettijohn KA, Radvansky GA. 2016a. Walking through doorways causes forgetting: event structure or updating disruption? *Q. J. Exp. Psychol.* 69:2119–29

Radvansky GA, Zacks JM. 2014. *Event Cognition*. New York: Oxford Univ. Press

Radvansky GA, Zwaan RA, Federico T, Franklin N. 1998. Retrieval from temporally organized situation models. *J. Exp. Psychol. Learn. Mem. Cogn.* 24:1224–37

Schwann, S., Hesse, F. w., & Garsoffky, B. (1998). The relationship between formal filmic means and the segmentation behavior of film viewers. *Journal ofBroadcasting* & *Electronic Media,* 42, 85-97.

Schwann, S., Garsoffky, B. & Hesse, F.W. Do film cuts facilitate the perceptual and cognitive organization of activity sequences?. *Memory & Cognition* 28, 214–223 (2000). https://doi.org/10.3758/BF03213801

Shimamura, A. P., Cohn-Sheehy, B. I., Pogue, B. L., & Shimamura, T. A. (2015). How attention is driven by film edits: A multimodal experience. *Psychology of Aesthetics, Creativity, and the Arts, 9*(4), 417–422. https://doi.org/10.1037/aca0000025

Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus 6:149-172.

Smith, Tim, J., (2012). The Attentional Theory of Cinematic Continuity. *Berghahn Journals* Volume 6, Issue 1, summer 2012: 1-27.

Sols, I. et al. (2017) Event boundaries trigger rapid memory reinstatement of the prior events to promote their representation in long-term memory. *Curr. Biol.* 27, 3499-3504.

Spiers, H.J. et al (2015) Place field repetition and purely local remapping in a multicompartment environment. *Cereb. Cortex* 25, 10-25.

Stewart, S. et al (2014) Boundary coding in the rat subiculum. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369, 20120514

Sun, C., Yang, W., Martin, J. *et al.* Hippocampal neurons represent events as transferable units of experience. *Nat Neurosci* 23, 651–663 (2020). https://doi.org/10.1038/s41593-020-0614-x

Swallow KM, Kemp JT, Candan Simsek A. 2018. The role of perspective in event segmentation. *Cognition* 177:249–62

Tang H, Singer J, Ison M, Pivazyan G, Romaine M, Frias R, Meller E, Boulin A, Carroll JD, Perron V, Dowcett S, Arlellano M, Kreiman G. (2016). Predicting episodic memory formation for movie events. *Scientific Reports*, 6:30175.

Tauzin T. 2015. Simple visual cues of event boundaries. *Acta Psychol.* 158:8–18

Tulving E, Markowitsch HJ (1998) Episodic and declarative memory: role of the hippocampus. Hippocampus 8: 198-204.

U. Rutishauser, E. M. Schuman, A N. Mamelak, Online Detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo. *J. Neurosci Methods* **154**, 204-224 (2006).

Whittington et al., The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation, Cell (2020), https://doi.org/10.1016/j.cell.2020.10.024

Woodward AL. 2009. Infants' grasp of others' intentions. *Curr. Dir. Psychol. Sci.* 18:53–57

Wilkenheiser AM, Reddish AD (2015) Hippocampal theta sequences reflect current goals. Nat Neurosci 18:289-294.

Zacks JM, Kurby CA, Eisenberg ML, Haroutunian N. 2011. Prediction error associated with the perceptual segmentation of naturalistic events. *J. Cogn. Neurosci.* 23:4057–66.

Zacks, Jeffrey M. *Flicker: Your Brain on Movies*. Oxford University Press, 2015.

Zacks, Jeffrey M. (2020). Event Perception and Memory. *Ann. Review of Psychol.* Vol 71:165-191.

Zacks, J.M. et al. (2001) Human brain activity time-locked to perceptual event boundaries. *Nat. Neurosci.* 4, 651-655.

Zacks, J.M. et. al. (2010) The brain's cutting-room floor: segmentation of narrative cinema. *Front. Hum. Neurosci.* 4, 1-15.

Zheng, J. Gomez Palacio Schjetnan, A, Taufik, V.A., Mamelak, A. N., Chung, J.M. Rutishauser, U., Kreiman, G. Event boundaries shape memory formation: evidence from single unit recordings in humans. Program No. 698.10. Neuroscience 2019 Abstracts. Chicago: Society for Neuroscience, 2019. Online.