Title
Neural Evidence of Mental Models in Movie Viewing: The Role of Narrative and Narrational Features

Author
Chen, Yibei

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Neural Evidence of Mental Models in Movie Viewing: The Role of Narrative and Narrational Features

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Commnunication

by

Yibei Chen

Committee in charge:

Professor René Weber, Chair

Professor Ronald E. Rice

Professor Hongbo Yu

June 2023

The dissertation of Yibei Chen is approved.

Ronald E. Rice

Hongbo Yu

René Weber, Committee Chair

April 2023

Neural Evidence of Mental Models in Movie Viewing: The Role of Narrative and Narrational Features

by

Yibei Chen

ACKNOWLEDGEMENTS

Every time I read someone else's dissertation; I always go to their acknowledgements page first. From the names and the words associated with those names, I can piece together the most important four to seven years of their life. Today, it is my chance to say THANK YOU to all the valuable people in the past four years of my life.

First, I want to thank my advisor and the PI of the Media Neuroscience Lab, Dr. René Weber. This lab was the reason I came to the U.S., and I am still glad that I made that choice four years ago. Thank you, René, for acknowledging my statistics/math/computational skills and leading me on the journey of neuroscience. Without you, this dissertation would never have happened. Without you, I would not have become this version of myself—a researcher who is dedicated and determined to neuroscience. Thank you, René, for the eye-opening experience I have had in this lab, which will be stamped indelibly in the rest of my academic career.

I want to say thank you to my committee members, Dr. Ron Rice and Dr. Hongbo Yu. Ron, thank you for always being patient, caring, and supportive. You have been the one who has seen the most of my tears and pain and the one who shares my passion for martial arts and fiction. Thank you for the endless support during my darkest times. Hongbo, as a mentor, as a friend (and even like an elder brother), thank you for valuing me and my research. I have had a happy and invaluable time at the YES lab. I also want to give special thanks to Dr. Greg Ashby, who taught my fMRI analysis and has been my statistics magician. Thank you, Greg, for always wanting the best for me. In addition, I want to thank Drs. Andrew Flanagin, Miriam Metzger, Andy Merrolla, Jennifer Kam, Karen Myers, and many other professors who have cared for and helped me with endless kindness.

I want to thank former and current MNLers who have helped me at different stages of my time here. Thanks to Richard Huskey, Chelsea Lonergan, Jacob Fisher, Frederic Hopp, Paula Wang, Musa Malik, Sungbin Youk, Kylie Woodman, and Brittany Wheeler. Thank you for being wonderful lab mates. I wish each one of you great success in the future. I also want to thank YES lab members, Amber Chen, Maddy Kashyap, Ava Ma de Sousa, Suyi Leong, Lu Zang, Jie Chen, and Mesude Okhan. All of you have made the lab experience so lovely. Thank you for giving me a second home.

I am very lucky to have the greatest friends in the world. Thank you, Jesse King, Laurent Wang, Zijian Lew, Li Qi, Chris Otmar, Nancy Molina-Rogers, Nitzan Navick, Allison Mazur, Carmen Chan, Ceciley Zhang, Qing Huang, Minghui Wang, Xingyu Liu, Lindsay Miller, Sandi Moxley, Abdullah Salehuddin, Nan Wilkenfeld, Kite Liu, Lauren Poole, Leslie Mavrakis, and Lillian Hughes. Each of you has lit a light in my life. Special thanks to Ruobing Han and Xian Feng who have been unconditionally supporting me from the other side of the Pacific Ocean for countless days and nights. Long Live Friendship! Thanks to my RAs, Tian Ding, Lori Lou, Adele Zhang, and Kaitlin O'Kane for their hard work.

A special shout-out to Sam Nastase. Thanks for the shared mind.

Lastly, and of course, most importantly, I express my deepest appreciation to my parents, who have biologically and intellectually raised me. Thank you for everything.

VITA OF YIBEI CHEN
April 2023

EDUCATION

Bachelor of Arts in Communication, Wuhan University, June 2014
Master of Arts in Communication, Fudan University, June 2019
Doctor of Philosophy in Communication, University of California, Santa Barbara, June 2023 (expected)

PUBLICATIONS

7. Chen, Y., Hopp, F.R., Malik, M., Wang, P.T., Woodman, K., Youk, S., & Weber, R. (accepted). Reproducing FSL's fMRI data analysis via Nipype: Relevance, challenges, and solutions. *Frontiers in Neuroimaging*.
6. Malik, M., Hopp, F. R., Chen, Y., & Weber, R. (2021). Does regional variation in pathogen prevalence predict the moralization of language in COVID-19 news?. *Journal of Language and Social Psychology*, 0261927X211044194.
5. Chen, Y., Sun, S., Zhao, X., Zhou, H., & Wang, F. (2020). Objective facts or misleading hype? Associations between features of e-cigarette marketing and sales on a Chinese e-commerce platform. *Int. J. Environ. Res. Public Health, 17*, 6711.
4. Sun, S., Zhai, Y., Shen, B., & Chen, Y. (2020). Newspaper coverage of artificial intelligence: A perspective of emerging technologies. *Telematics and Informatics*, 101433.
3. Sun, S., He, J., Shen, B., Fan, X., Chen, Y., & Yang, X. (2020). Obesity as a “self-regulated epidemic”: Coverage of obesity in Chinese newspapers. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity,* 1-16.
2. Sun, S., He, J., Fan, X., Chen, Y., & Lu, X. (2019). Chinese media coverage of eating disorders: Disorder representations and patient profiles. *International Journal of Eating Disorders*,1-10.
1. Sun, S., & Chen, Y. (2018). The disciplinary trajectory and topical landscape: Health communication research in mainland China from 1987 to 2016. *Journalism Bimonthly, 3*, 011.

AWARDS

George D. McCune Dissertation Fellowship, University of California, Santa Barbara, 2022

Abstract

Neural Evidence of Mental Models in Movie Viewing: The Role of Narrative and Narrational Features

by

Yibei Chen

The impact of narratives is evident in various aspects of society, from literature and historical documents to scientific explanations, political speeches, and everyday conversations. Research has shown that how people interpret narratives can influence cognitive and linguistic behavior. This dissertation focuses on mental models constructed during movie-watching, a particular type of narrative experience. By analyzing multiple existing datasets and designing three studies, this dissertation offers a framework for investigating the role of narrational features in mental model constructions in the brain across three levels (micro, macro, and super) during movie viewing.

The first study examines the contributions of narrative and narrational features to macro-level mental model construction in the brain. We found a significant relationship between narrational features and low-level brain regions but not between narrative features and high-level brain regions in the event segmentation results. These low-level regions may not be as sensitive to the specific details of the narrative features as they are to how the story is presented through the narrational features. The narrational features may provide cues or signals to the high-level brain regions, guiding their narrative interpretation. This is consistent with previous research that

has found that how a story is told can significantly impact how it is perceived and remembered by audiences.

The second study investigates how the brain reacts to narrative features, particularly moral-relevant content, while constructing micro-level mental models. We find consistent results across model that macro-events with more micro-events are more likely to have brain-data boundaries overlapped with human-annotated boundaries. However, our generalized linear models did not find evidence to support our hypothesis that higher levels of moral-relevant content would correspond to higher inter-subject correlations, indicating engagement.

The third study investigates how the brain uses narrative and narrational features in super-level mental model construction, specifically in schema maintenance and violation. Our results showed that intra- and inter-subject correlations in the precuneus were significantly higher for the intact clip than for the scrambled clip, indicating the precuneus's involvement in schematic thinking during narrative processing, particularly in posterior medial regions. However, we did not find a significant relationship between inter-subject correlation (i.e., engagement) and the deviation of segments (broken schemas).

Researchers in communication and media psychology can learn from the neurological component of this dissertation since it offers a biological perspective and methodological innovations to advance our understanding of narration and narrative effects in the brain. Next, the operationalization of narrative features and links between features or combinations of features and brain activity can be an exemplar for neuroscientists less familiar with media studies.

TABLE OF CONTENTS

GLOSSARY

Cofluctuation: The simultaneous fluctuation of two or more variables, typically referring to brain activity patterns that change together in response to a stimulus or task. See Time-resolved ISC.

Event Segmentation: The cognitive process through which people divide continuous experiences into discrete events or meaningful units, allowing for easier encoding, storage, and retrieval of information in memory.

HMM (Hidden Markov Model): Statistical models used for modeling time series data or sequences, in which the system being modeled is assumed to be a Markov process with unobservable or hidden states. In the context of event segmentation, HMMs can be employed to identify hidden states representing different events or cognitive processes underlying observed neural or behavioral data.

GLM (Generalized Linear Model): A flexible statistical model that extends linear regression to accommodate a wider range of response variables by incorporating different probability distributions and link functions.

Inter-subject Correlation: A measure of the similarity between the neural responses of different individuals when exposed to the same stimulus or performing the same task, indicating shared cognitive or perceptual processes.

Intra-subject Correlation: A measure of the consistency of an individual's neural responses across multiple exposures to the same stimulus or repetitions of the same task, indicating stable cognitive or perceptual processes.

Mental Models: Internal cognitive representations of external reality, which allow individuals to understand, predict, and interact with the world around them.

Macro-level Mental Models: Mental models that represent higher-order, abstract concepts or structures, which help individuals understand and make sense of broader systems or situations.

Micro-level Mental Models: Mental models that represent specific, detailed information about particular objects, events, or relationships, which help individuals navigate and make decisions within specific contexts.

Schemas: Cognitive structures or frameworks organize and guide the processing, storage, and retrieval of information, based on prior knowledge and experience.

Super-level Mental Models: Mental models that encompass overarching principles or frameworks for understanding and organizing multiple macro- and micro-level mental models. Schemas are Super-level mental models can be considered as a type of schema that provides a higher-order structure for integrating and organizing more specific mental models.

Time-resolved ISC: Time-resolved Inter-subject Correlation (ISC) is a method for analyzing inter-subject correlations on a finer temporal scale, which can reveal the degree of cofluctuation in neural responses between individuals at specific time points or intervals during stimulus presentation or task performance.

Introduction

Narratives are pervasive in various aspects of society, from literature and historical documents to scientific explanations, political speeches, and everyday conversations (Nash, 1994). Research has shown that how people interpret narratives can impact cognitive and linguistic behavior (Gerrig, 1993; Pennebaker & Seagal, 1999). By studying how individuals comprehend narratives in the brain, we can learn more about how the brain processes information (Armstrong, 2020), how people acquire and organize knowledge, and how they evaluate themselves and the world around them (Gergen & Gergen, 1988), and how they construct their sense of reality (Bruner, 1991). Examining how cultural codes and norms influence individuals in their narrative interpretation can provide insight into the relationship between culture and cognition (Rasmussen, 1999). This dissertation focuses on the mental models individuals construct during movie-watching, a particular narrative experience.

Comprehenders[1] use mental models to organize, interpret, and remember the events and information in the narrative. Mental models (i.e., events[2]) are structures in memory that represent the key elements and relationships between them in the narrative. Mental models play an essential role in narrative comprehension. Within each mental model, the strength of how key elements and their relationships in the narrative are represented can be measured through collective engagement across the audience (i.e., inter-subject correlation). More robust engagement can facilitate narrative comprehension (Regev et al., 2013). These mental models are dynamic and are continually updated as new information is encountered in the narrative (van

[1] In this paper, *comprehenders* refer to readers (of books, newspapers, magazines, etc.), audience (of movies, TV shows, concerts, etc.), and all other types of narrative consumers.
[2] Broadly speaking, events are part of mental models. According to Event Segment Theory (Zacks et al., 2009), in the context of narrative, mental models are events. In this dissertation, we will use mental models, events, mental events, interchangeably.

den Broek et al., 1999). The construction of mental models is a fundamental aspect of narrative comprehension, and understanding the process of mental model construction can provide insights into how individuals comprehend and remember narratives. Furthermore, mental models are constructed during narrative comprehension and influence the comprehension process by guiding attention and interpretation (Zwaan et al., 1995). Therefore, investigating the construction of mental models is crucial for understanding how individuals comprehend and are influenced by narratives in various forms of communication.

Mental models refer to mentally constructing what has happened in the narrative through symbolic perception, inference, and reasoning, which incorporate information from the narrative itself along with personal experience and knowledge (Barsalou, 1999; Bower & Morrow, 1990; Gernsbacher et al., 2004; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). Research has pointed out that mental modal construction can happen at the macro (i.e., across events), micro (i.e., within events), and super (i.e., schemata) levels (e.g., Zacks & Swallow, 2007; Zwaan, 2016). However, current literature in communication or neuroscience has rarely examined those three levels simultaneously. The current dissertation fills this gap by designing three studies to explore how the brain processes narratives and constructs mental models at each of those three levels and how narrative and narrational features play a role in these processes.

Below, we first provide an overview of mental model construction at macro, micro, and super levels, explain what narrative and narrational features are and why they play an essential role in building mental models, and review current studies on this topic. This dissertation conducts three studies for each level accordingly. In each study, we address research questions and hypotheses, describe the dataset(s) in use, provide detailed descriptions of the analytical approach, present the results, and discuss findings and limitations. Last, we provide a short general discussion of the whole dissertation.

Literature Review

As previously stated, narratives can encompass various formats, including text, audio, video, and others. However, this dissertation focuses explicitly on narratives presented through movies. Therefore, "narrative" in the singular form pertains to the storyworld within the movie and should not be confused with "narration," which pertains to how the story is presented through its chosen medium.

Narrative and narration are essential components in understanding and constructing mental models of a storyworld, particularly in movie viewing. The narrative represents the story's content, including the characters, actions, time, and places. At the same time, the narration pertains to how the story is presented or conveyed to the audience. While the narrative mainly determines the content of mental models, the narration influences the ease or difficulty of constructing accurate mental models. Narrative features are considered high-level features, including the story's who, what, where, and when. In contrast, narrational features, such as pitches, color, and luminance, are low-level features that do not necessarily contribute to the narrative but rather the physical aspects of the movie itself. Movies are intentionally designed to manage viewers' attention, directing them to perceptible features responsible for the film's affective, depictive, narrative, and semantic content. Filmmakers use various techniques such as camera movements, lens movements, and editing techniques to create and manipulate narrational features, ultimately influencing the audience's understanding and mental model construction of the story.

Movie viewing, like other types of narrative comprehension, involves the dynamic process of constructing mental models (Freyd, 1987; Gernsbacher et al., 2004; Johnson-Laird, 1983). Both narrative and narration play crucial roles in this process, with the narrative essentially determining what goes into the mental model and the narration impacting the ease or difficulty of constructing an accurate mental model. Mental model construction occurs at three levels: micro, macro, and super. The micro level pertains to what happens within each event (i.e., locally), the macro level concerns the connections between these events (i.e., globally), and the super level encompasses the schemas used in the narrative. Herman's (2004) integration of narrative theory with linguistics and cognitive science posits that a story world is a sequence of events, with microdesigns referring to the local states that contain the verbal (words), mental (emotions), and physical (actions) states of characters and a consecutive series of events, while macrodesigns encompass the temporal, spatial, and causal structures of the entire story.

As mentioned, schemas are an essential component of mental model construction at the super-level. A schema is a knowledge structure that binds the information in memory and can be triggered by narrative cues (e.g., environment settings in a movie) to facilitate comprehension (P. Whitney et al., 1995). However, the super-level does not dominate the micro and macro levels in a hierarchical structure. Schemas can also be used at both the micro and macro levels. For example, at the micro-level, we may anticipate seeing handcuffs and cells, a sheriff or marshal, and other deputies when a police station appears on the screen. At the macro level, we may expect the first identified suspect in detective fiction to be innocent at the end of the story. Regardless of the level, narrative and narrational features are crucial in constructing mental models.

Micro-level Features

At the micro-level of mental model construction, each event in the narrative is represented. Key narrative features at this level include who is involved in the event, the main protagonists and supporting characters, the actions they take, and the mental states they experience. Protagonists are central to the story while supporting characters contribute to the narrative arc of the main characters. Every action has consequences, which can result from previous events or be the catalyst for future developments. Emotional states are also crucial at this level and play a significant role in the narrative process. Emotional expressions, such as facial or verbal cues, can help viewers empathize with the characters and become more absorbed in the narrative. Research has shown that emotional engagement with characters can enhance comprehension and overall enjoyment of the story (Mar et al., 2011; Murphy et al., 2013; Rooney & Hennessy, 2013).

Emotions can also provide clues for updating existing beliefs or predicting upcoming events. This is because emotions are reactions to what has happened and have action potential (Parrott, 2002). For instance, if a character's emotion contradicts our general knowledge (e.g., a man appears very calm when he finds out that his wife has cheated on him), we can update our previous inference (e.g., we may infer that the man is not as affected by the news as we had initially thought) or generate expectations regarding upcoming events (e.g., we may anticipate that the man is planning an act of revenge and is merely suppressing his anger).

At the micro-level, narrational features work in tandem with narrative features to enhance emotional engagement with the story world. Large images, achieved through techniques such as zoom-ins, can increase attention and lead to heightened arousal and emotional engagement. In genres such as action, suspense, or horror, the use of looming objects on the screen can contribute to tension building and cause the audience to feel as anxious or scared as the characters. Color choices in movies can also offer valuable cues to characters' emotions, with cool and dark colors associated with sadness and misfortune. In contrast, warm and bright colors indicate happiness and hope. Audio features, such as the movie's soundtrack, also play an essential role in emotional engagement, as different pitches and tones can lead to varying degrees of emotional arousal. Research has shown that a well-crafted soundtrack can elicit an emotional response and influence the audience's interpretation of the story (Barbas et al., 2011; Grosso et al., 2015).

Marco-level Features

The macro-level narrative features focus on grouping information into events and specifying event boundaries. This cognitive strategy increases processing speed in the brain by automatically grouping information into scenes during movie viewing (Zacks et al., 2009). Temporal and spatial information help define event boundaries, with spatial information being a more robust indicator than temporal information. Changes in the protagonist's location often coincide with changes in time, making it easier to identify event boundaries. For example, in the movie *Source Code* (2011), events can be segmented by time - every time the protagonist wakes up - since the location (i.e., train) remains constant. Alternatively, in the movie *Prometheus* (2012), event boundaries can be set based on location - inside the spacecraft or the cave (i.e., the alien spaceship).

Spatial information is often a more reliable cue for identifying event boundaries than temporal information. This is because location changes are typically associated with changes in time, and certain actions can only occur in specific locations. When temporal and spatial changes coincide, it is usually easy to determine where one event ends and another begins.

Information about the characters can also assist in segmenting events. The main character(s) typically interacts with different characters in different scenes, cueing scene transitions. Even when interacting with the same group of characters, variations in their emotions during interactions across scenes can also signal a change in the event.

Regarding narrational features, some movies provide direct information about time and space changes through titles or other explicit segment signs. For example, *The Grand Budapest Hotel* (2014) uses titles (e.g., Part 1: M. Gustave) to indicate that everything until the next Part belongs to the same theme. However, for movies without explicit segment signs, features such as luminance and color can help identify event boundaries (Cutting et al., 2010). Luminance, which measures light intensity in an image or series of images, can be manipulated during shooting and postproduction to elicit specific effects. For instance, bright light and high luminance can create a sense of other-worldliness. Dreams or memories usually have different luminance than the current timeline, except in movies like *Inception* (2010), where the distinction between dreams and reality is intentionally blurred.

Color is also a valuable tool for identifying changes in time and space. Hue and saturation are standard parameters used to measure color in various spaces. Hue refers to the color itself, such as red, blue, and green, while saturation refers to the intensity of the color. Color changes, particularly in hue, can indicate changes in scenes (Cutting et al., 2011), as time, space or content shifts are often associated with color changes.

Super-level Features

At the super-level of narrative features, schemas play a crucial role in shaping the audience's interpretation of the story. Schemas are pre-existing mental frameworks that help us make sense of the world around us. They influence our expectations of events and guide our attention toward schema-relevant information while inhibiting other possible inferences (Schank & Abelson, 2013). Once established, schemas strongly influence how the audience interprets characters' intentions, behaviors, and event sequences (Ghosh & Gilboa, 2014). Movies often use ambiguous materials that plausibly match but contradict schemas to create suspense, novelty, and surprises. Irrelevant or unimportant information is also often ignored to save cognitive resources.

In addition to their cognitive and psychological benefits, schemas also have social benefits in facilitating successful communication. When individuals adopt similar schemas, they demonstrate a shared understanding of the external world, which is crucial for effective communication. A study by Lahnakoski et al. (2014) found that participants who shared similar schemas during movie watching showed higher synchronous brain activity than those with different schemas. Participants who shared the detective schema better recalled specific events or characters in the video, while those who shared the decorator schema reported more details about the interiors and yards. Filmmakers intentionally manipulate schemas to achieve particular effects, and the audience consciously or subconsciously interprets them to facilitate understanding.

Schemas often operate abstractly and can be repeatedly applied across narratives. It is not the particular narrative feature (such as time, place, or character) itself but the combination of features (such as the story, setting, or environment) that triggers schematic thinking. Filmmakers intentionally manipulate these combinations to achieve particular effects, and the audience consciously or subconsciously interprets them to facilitate understanding. For example, a knife in the kitchen in a romance movie may not attract much attention, but a knife in a detective movie can immediately trigger specific imaginations. Narrational features, such as shot structures and transitions, also contribute to schematic thinking by constructing connections among separate elements within or across events. Crosscutting is a commonly used transition technique that changes back and forth between scenes so that actions occurring in different locations seem to be unfolding at the same moment. For example, if the first shot is about a shooter preparing for a long-range shooting and the next shot is a person jogging, the audience will likely infer that the jogging person is the shooter's target.

Narrative Processing in the Brain

There is a growing body of research in cognitive neuroscience that aims to understand how the brain processes narrative and creates mental models in response to naturalistic stimuli such as audiobooks, music, and movies. Compared to static and highly controlled stimuli, naturalistic stimuli have the advantage of being able to replicate real-life scenarios and evoke brain responses that are more ecologically valid (Sonkusare et al., 2019). Among naturalistic stimuli, movies are instrumental as they are intentionally structured, designed, and edited to capture and maintain audience engagement (Jääskeläinen et al., 2021; Vanderwal et al., 2019). Therefore, movies can be practical tools for studying how the brain infers meaning and constructs mental models.

As Finn et al. (2022) outlined, using naturalistic stimuli in brain research has led to several topics, three closely related to narrative comprehension. These three topics include the hierarchical nature of brain structure, brain network and dynamics, and memory processing, specifically encoding and retrieval.

The research on brain hierarchies with naturalistic stimuli focuses on the differences between low-level and high-level brain areas while encoding such stimuli. Sensory information processing areas, such as the primary visual cortex and motion-sensitive area MT+, have high response reliability regardless of the content. In contrast, high-level brain areas responsible for comprehension and reasoning, such as the precuneus, superior temporal sulcus (STS), posterior lateral sulcus (LS), and temporal-parietal junction (TPJ), have context-dependent and time-consuming responses. This indicates that different regions of the brain respond differently to narrative and narrational features. These findings were reported by Brennan, Ren, and Hasson (2016), Brennan and Hale (2019), and Hasson et al. (2008).

Research on brain networks and dynamics has focused on how the brain processes narratives in terms of their spatial and temporal dynamics (Andric et al., 2016; Chai et al., 2016; Jang et al., 2017; Meer et al., 2020; Ogawa, 2021; C. B. Young et al., 2017). For instance, researchers have found that brain network configurations change over time when repeatedly watching the same movie, while local activity profile for each brain region (i.e., nodes in the brain network) are stable over time during one-time movie viewing (Andric et al., 2016; Meer et al., 2020). The brain network dynamics in response to different cognitive demands evoked by movie features suggest that movie viewing involves a relatively fixed set of brain regions, but the connection dynamics within and across regions and individuals vary based on the cognitive demands elicited by narration (Jang et al., 2017; C. B. Young et al., 2017).

Finally, research on memory has focused on the encoding and retrieval of information during narrative processing (Aly et al., 2018; Masís-Obando et al., 2022). Schemas, events in a specific sequence, are stored in our memory system. When the sequence of events is scrambled, it can disrupt the functional connectivity among or within brain regions responsible for schematic thinking. Aly et al. (2018) found that posterior medial regions demonstrated reliable temporal dynamics when the movie content was consistent with schemas but not when it was inconsistent. This line of research suggests that encoding and retrieving schemas during narrative processing, such as during movie viewing, can contribute to the overall understanding of the story.

In addition, new analytical techniques have been developed and refined to facilitate naturalistic neuroscience research better. These methods include event segmentation algorithms (e.g., Hidden Markov Modeling and Greedy State Boundary Search) and time-resolved inter-subject correlation analysis (Esfahlani et al., 2022; Liu et al., 2022; Tanner et al., 2022; Wass et al., 2019).

Event segmentation algorithms are methodological tools that extend the "Event Segment Theory" (Zacks & Swallow, 2007) and consider the hierarchy across brain areas and differences in stimuli over time. Event segmentation is a cognitive process of perceiving and organizing ongoing experiences into meaningful units or events. According to Event Segment Theory, events are defined as relatively discrete and temporally extended episodes that have a beginning and an end, and are perceived as having a coherent structure and meaningful purpose. Applying these methods to naturalistic stimuli can provide insight into how the brain perceives, encodes, and retrieves information during comprehension. Event segmentation can be used to assess mental model (mental event) construction at both the macro and micro levels.

The time-resolved inter-subject correlation analysis explores the moment-to-moment co-fluctuations of neural activity and reconstructs dynamic functional coupling patterns across participants. Researchers can link narrative and narrational features with brain activity as the stimuli unfold in real time by tracking these co-fluctuations. For example, within each macro or micro mental event, the time-resolved inter-subject correlation can demonstrate which narrative features evoke higher engagement, which further facilitates narrative comprehension. These methods represent promising approaches to the naturalistic neuroscience field, allowing for a more detailed and comprehensive understanding of brain activity during complex, real-world experiences. In this dissertation, we use the time-resolved inter-subject correlation to examine how engaging each mental model (i.e., event) will be.

The Current Dissertation

This dissertation investigates the three-level mental model construction during movie viewing and aims to make three contributions to communication, media psychology, neuroscience, and media production.

First, the neural aspect of this work can provide insights for communication and media psychology scholars interested in narrative effects and how and why specific narratives are powerful. This dissertation will provide neural evidence of the three-level mental model construction and promote our understanding of how deliberately crafted, multimodal, dynamic stimuli, such as media products like films, influence audiences via brain processing at these three levels. While communication and media scholars focus on behavioral and psychological aspects of narrative processing, cognitive neuroscience provides a biological perspective. It brings methodological innovations to advance our understanding of narrative effects and other communication phenomena.

Second, this dissertation can also benefit the neuroscience community by linking narrative and narrational features with brain activity and approaching narrative processing from three layers rather than one piece. Media stimuli generate rich cognition, emotions, and behaviors but have long been ignored in neuroscience research due to their complexity. This dissertation incorporates various narrative and narrational features. It extensively explores the relationships between different features or combinations of features and brain activities, which will push the envelope of studying narratives in neuroscience.

Third and last, media practitioners, such as filmmakers and screenwriters, intend to achieve specific goals with their media products, artistically and/or financially. The three-level mental model construction can offer filmmakers insights, endorsed by neural evidence, on how the audience generally processes movies. Those insights, incorporated with filmmakers' practical experience, can help to inform media production. Moreover, this dissertation explores narrative and narrational features, which filmmakers and screenwriters have painstakingly designed to achieve the intended effect. On this note, this dissertation's explorative analyses can enhance media practitioners' understanding of how their deliberately crafted macro- and micro-designs synergically affect the audience's brain.

To achieve these contributions, this dissertation analyzes multiple existing datasets, designs three studies, and provides a framework to investigate the role of narrative and narrational features in mental model constructions across three levels (micro, macro, and super) during movie viewing. The first study examines the contributions of narrative and narrational features to macro-level mental model construction in the brain. The second study investigates how the brain reacts toward narrative features (especially moral-relevant content) when processing micro-level mental models. The third study explores how the brain utilizes narrative and narrational features in super-level schematic thinking. By conducting these studies, this dissertation will provide empirical evidence of how narrative and narrational features affect the construction of mental models during movie viewing at different levels, from the micro to the macro to the super level. The results of these studies will contribute to the fields of communication and media psychology, cognitive neuroscience, and media production.

In summary, this dissertation aims to investigate the three-level mental model construction during movie viewing and contribute to understanding the neural, cognitive, and practical aspects of narrative processing. This work can offer insights for scholars in communication and media psychology, cognitive neuroscience, and media production by exploring the relationships between narrative and narrational features and brain activity at different levels. The findings of this dissertation can advance our knowledge of how media products exert influence on audiences via brain processing and can inform media practitioners on how to craft narratives that achieve their intended effects.

Study 1 Macro-level Mental Model Construction

Research Questions and Hypotheses

In this study, the aim is to understand how the brain processes macro-level mental models during movie viewing by analyzing the contributions of narrative and narrational features. Macro-level mental models are major events or groupings of similar information in a movie that summarize the meaning of multiple micro and small events. For instance, in a romance movie, major events could include the characters' first meeting, developing feelings toward each other, falling in love, having conflicts, and eventually breaking up or having a happy ending. The brain can integrate pieces of information into coherent events or ideas, and macro-level mental models serve as high-level sensemaking of the movie's narrative. This study will explore how the narrative and narrational features influence macro-level mental model construction in the brain during movie viewing.

The film *500 Days of Summer* (2009) shows how filmmakers can deliberately signal critical points to the audience. Using a non-chronological-ordered calendar, the movie explicitly conveys the main characters' love affair's essential moments. Each time the "Day" calendar appears, it prompts the audience's brain to construct a narrative segment and facilitates understanding of the movie at the macro level. Micro-events occur every two "Days." There are boundaries between macro-events and micro-events within this hierarchically nested mental model construction. These boundaries across macro-events should be larger than those across micro-events in the same macro-event.

In event segmentation theory, event boundaries are perceived when further information becomes substantially less predictable due to changes in narrational or narrative (i.e., contextual) cues, such as color, motion, shots, characters, goals, and spatial location (Zacks, 2020). Thus, both narrative and narrational features contribute to event segmentation during narrative comprehension. Moreover, a well-designed movie should ensure that the narration and narrative serve each other, such as changes in characters' moods associated with background music changes. This type of connection between narrative and narrational features occurs at the macro and micro levels, so boundaries between major events in the narrative should also serve as boundaries in the narrational features.

Based on this, we propose that there is a significant relationship between event boundaries in the narrative and in the narrational features.

H1: Event boundaries in narrative and in narrational features are significantly related.

Furthermore, recent neuroscience studies (Baldassano et al., 2017; Geerligs et al., 2021) have developed techniques to segment events based on brain activity during movie watching and have identified a hierarchical organization in event segmentation across the cortex (Geerligs et al., 2022). The lower-level sensory regions, such as the visual, auditory, and somatosensory cortices, have shorter events, while higher-level regions, such as the medial prefrontal gyrus and anterior portions of the lateral prefrontal cortex, have longer events. These findings suggest that event segmentation in different brain areas relies on different features of the stimuli. Geerligs et al.'s (2022) research utilized an 8-minute black-and-white movie clip and the Greedy State Boundary Search (GSBS) algorithm to segment events. However, Study 1 aims to replicate these findings on a full 90-minute movie with richer audiovisual information, utilizing the Hidden Markov Model (HMM) approach, instead of the GSBS algorithm. Based on my previous experience, GSBS has been found to work well on short time series but has difficulties with longer time series, which is why the HMM approach was chosen further explained below.

H2: The organization of neural event boundaries is hierarchical in the temporal cortex, with low-level regions (such as the auditory cortex and visual cortex) having (a) more and (b) shorter events and high-level regions (such as the medial prefrontal cortex and precuneus) having (c) fewer and (d) longer events.

Scholars have found that the reproducibility of brain responses differs across different brain regions. While low-level brain areas can form reproducible response patterns regardless of temporal disruptions, the reliability of responses in several higher brain areas is affected by information accumulated over longer time scales (Hasson, Yang, et al., 2008). Furthermore, research has shown that highly reproducible fMRI responses are primarily attributed to the high-level natural content in a movie. In contrast, low-level visual features without actual content (e.g., no characters or objects) in a scrambled movie evoke significantly reduced degrees and extent of reproducible responses (Lu et al., 2016). These findings suggest that narrational and narrative features are processed within different regions and scales of the brain and that the processing of narrational features is reproducible when narrative features are present. Therefore, when watching a movie with both narrational and narrative features, we expect low-level brain regions to utilize narrational features more than high-level regions do. Conversely, high-level brain regions are expected to utilize narrative features more than low-level regions.

H3: Event boundaries in narrational features are more related to event boundaries in low-level than high-level brain regions.

H4: Event boundaries in narrative features are more related to event boundaries in high-level than low-level brain regions.

In this study, we investigate the neural mechanisms underlying macro-level mental model construction during narrative processing. Specifically, we aim to examine how the brain processes narrational and narrative features and how these types of features contribute to forming event boundaries in the viewer's mental model of the movie. To achieve this goal, we analyze fMRI data collected. At the same time, participants watched the feature film *500 Days of Summer* and also used movie annotation data to identify the narrational and narrative features that contribute to the construction of the viewer's mental model. We test four hypotheses related to the processing of narrational and narrative features and their contributions to event boundary formation in the viewer's mental model. If supported, our hypotheses would provide insights into the neural basis of macro-level mental model construction and advance our understanding of the cognitive and neural processes involved in narrative comprehension.

Materials and Methods

fMRI Data

Dataset description. This study utilizes fMRI data from the Naturalistic Narrative Database (Aliko et al., 2020), focusing on the feature film *500 Days of Summer*. The dataset includes 20 participants who watched the entire 90-minute film during scanning. Two participants were excluded, one because they were scanned with a different head coil and the other because they were only offered glasses after the first run. This left a total of 18 participants (9 male) for analysis. None of the participants had previously seen the film. The functional and anatomical images were acquired using a 1.5 T Siemens MAGNETOM Avanto with a 32-channel head coil, with a TR[3] of 1s. For more details on data acquisition and preprocessing, please refer to Aliko et al. (2020). The original data are available on OpenNeuro[4], and we use the

[3] Repetition time, the amount of time between successive pulse sequences applied to the same slice during MRI scanning.
[4] https://openneuro.org/datasets/ds002837/versions/2.0.0

preprocessed (i.e., standardized) data from (de la Vega et al., 2022) using fMRIPrep (Esteban et al., 2019).

To ensure data consistency and accuracy, we preprocessed the fMRI data and movie annotation data. Specifically, we removed the opening credits occurring before 00:00:37 and ending credits after 01:26:51. The corresponding fMRI volumes were also excluded from further analysis. Additionally, we applied a 4-TR delay in hemodynamic response to shift the fMRI data, following previous studies (Rajapakse et al., 1998).

Regions of interest (ROIs). ROIs were identified using Schaefer et al.'s (2018) 400-parcel 17-network parcellation. We selected two low-level regions, the auditory cortex and the primary visual cortex, and two high-level regions, the medial prefrontal cortex and the precuneus, based on their established role in narrative processing (Andric et al., 2016; Geerligs et al., 2022; Hasson, Yang, et al., 2008; Song et al., 2021). To obtain time series data, we extracted and standardized the data at the voxel level for each participant before averaging it into the parcel level. Finally, the time series data for each brain region were averaged across participants. To ensure data quality, the extracted time series were visually inspected for each participant.

Movie Annotations at the Macro Level

Narrative features. As mentioned, *500 Days of Summer* provides calendars ($N = 41$) about the love affair between two main characters. This explicit temporal information and visual disruption (the same picture with different numbers appears on the full screen) inevitably impact participants' narrative comprehension process. In this study, we use those 40 pre-defined boundaries[5] and identify the time point of each as the narrative boundary.

[5] The number of boundaries is equal to the number of events minus 1. For example, one boundary is needed to divide one event into two. Two boundaries are needed to divide one event into three.

In this study, we utilize the movie *500 Days of Summer*, which includes explicit temporal information in the form of calendars and visual disruptions that can impact participants' narrative comprehension processes. We identify the pre-defined boundaries from these calendars as narrative boundaries. This particular case of narrative feature construction highlights the importance of temporal, spatial, and character information changes, which provide cues for segmentation. Compared to other methods, using a movie with pre-defined narrative features allows us to use them as a natural ground truth without additional justification[6].

Narrational features. We extract low-level sensory and perceptual features from the movie 500 Days of Summer to investigate their role in narrative processing. Visual features include the level of abstractness, blurriness, brightness, daylight, indoor/outdoor scenarios, landscape, sharpness, and vibrance of a frame[7]. Audial features[8] include the constant-q chromatogram[9], Mel-frequency cepstrum[10], root mean square[11], and tonal centroids[12]. These features are time series with the same length as the movie and sampled at each second. While there are more visual features than audial features, some of the audial features have multiple dimensions, making the number of visual and audial features similar. To avoid potential confounding effects, we model visual and audial features separately. All features are extracted from Neuroscout[13] and truncated to match the fMRI data.

[6] In a more general approach (for other movies), we can annotate temporal changes when there is a significant shift in time and spatial changes when there is a shift in location. Additionally, we can quantify psychological and physiological traits for changes in character information, including emotional changes associated with a character's face, words, and voice.
[7] A video is a collection of frames across time.
[8] Audial features listed here are widely used in signal processing.
[9] The transformation of a sound to a frequency domain
[10] The representation of the short-term power spectrum of a sound
[11] The average loudness of an audio track within a given period (default is 300 milliseconds)
[12] The representation of the tonal space in audio
[13] https://neuroscout.org/predictors

Statistical Analysis

In this study, our approach to event segmentation involves using Hidden Markov Models (HMMs). HMMs are a statistical modeling technique that allows for modeling an underlying "hidden" state of the brain at each time point based on observed time series data (e.g., fMRI signals, narrational features). In this context, the hidden state can represent a particular mental or neural process, such as an event, corresponding to constructing a macro-level mental model. Neuroscience researchers have applied HMMs to naturalistic data and found that they can effectively detect event boundaries that are in close agreement with human annotations (Baldassano et al., 2017). Our input data for the HMMs consist of 2D time series (either n_features or n_parcels by n_timepoints) with a pre-defined number of events, and the output of the HMMs are the time points (or boundaries) of each event and the associated probability of each event.

To test H1 (the association between event boundaries in the narrative and those in the visual and audial narrational features), we use the calendar days (N=41) provided in *500 Days of Summer* as a proxy for macro-events in the narrative. We input the macro-events into the visual and audial HMMs separately, resulting in a list of time points as event boundaries for each. We then compare these boundaries to the list of time points for macro-events in the narrative.

To quantify the association between boundaries in the visual and audial narrational features and those in the narrative, we developed a measurement to compare the closeness between the two types of boundaries. Specifically, for a time point A in the output of the visual or audial HMM, we identify the closest time point B in the narrative boundaries. We then count A as a hit if the difference between A and B is smaller than a given threshold. The threshold value can range from as small as 0 (the most conservative approach) to as large as the length of the shortest event (the most liberal choice in a logical space). The hit rate, which is calculated by dividing the number of hits by the number of events, reflects the strength of the association between event boundaries in the visual and audial narrational features and those in the narrative.

To test the significance of the association between event boundaries in the visual and audial narrational features and those in the narrative, we used a non-parametric permutation test. First, we defined a hit rate measurement to quantify the similarity between the two sets of boundaries. For a time point in the visual or audial HMM outputs, we identified the closest time point in the narrative boundaries. If the difference between the two time points was smaller than a given threshold, we counted it as a hit.

Next, we randomly generated the same number of time points as in the actual data and calculated the hit rate between the visual or audial boundaries and the pseudo boundaries. We repeated this procedure 10,000 times to obtain a null distribution of the hit rate. The p-value of the null hypothesis (i.e., the real hit rate is not greater than zero) was calculated as the probability of the actual hit rate being smaller than the pseudo hit rates.

The above methods were applied separately to the visual and audial narrational boundaries.

To test H2 (the temporal cortical hierarchy in neural event boundaries), we first identify the optimal number of events in HMM for each of the low-level and high-level brain regions separately. The algorithm uses log-likelihood to assess the model's performance. The model with the largest log-likelihood is considered the best model, and the optimal number of events is the output from the best model. To obtain the log-likelihood, we split the data into training and testing sets. We run HMM on the training set to obtain a set of parameters and then fit the testing set into the same model. The test log-likelihood measures how well the training model fits the testing data.

To determine the optimal number of events for each brain region in testing H2, we perform HMMs with different inputs and compare the resulting test log-likelihood values. However, as the number of events can range from 1 to the total number of time points, running HMMs for every possible number of events is too computationally expensive. Therefore, we have developed an efficient iterative algorithm to identify the optimal number of events for each brain region. The algorithm starts with a short list of event numbers, uses it to fit the training and testing data, identifies the event number with the largest log-likelihood, and generates a new list of event numbers using this value as a seed. This process is repeated until the log-likelihood reaches saturation (i.e., the change in values converges to 0) and the optimal number of events is identified as the one with the largest log-likelihood across all iterations. By using this algorithm, we only need to fit an average of 120 HMMs instead of 5,174 for each brain region. Our focus is on the low-level regions of the auditory and primary visual cortex, and the high-level regions of the medial prefrontal cortex and the precuneus. Additionally, we test this algorithm on all 17 networks for comparison.

To test H3 (whether event boundaries in narrational features are associated more with boundaries in low-level than high-level brain regions), we adopt a similar approach as in testing H1. We use the optimal number of events in each brain region (i.e., the auditory cortex, the primary visual cortex, the medial prefrontal cortex, and the precuneus) obtained from H2 to fit different HMMs for visual and audial narrational features. Then, we identify the significance of the hit rate by comparing the permutated ($N = 10{,}000$) hit rate of narrational features with low-level brain regions and that of narrational features with high-level brain regions.

To test H4 (whether event boundaries in the narrative are associated with boundaries in high-level brain regions), we adopt a similar logic but in a slightly different manner. We fit HMMs in four different brain regions using the number of macro-events in the narrative (N = 41), and then calculate the permutated (N = 10,000) hit rate of boundaries from each of those brain HMMs to boundaries of macro-events in the narrative.

Results

H1. As shown in **Table 1.1** and **Figure 1.1a**, H1 is supported. The hit rate between event boundaries from narrational features and those from the narrative was calculated with 10,000 permutations, and significant hit rates were found. The hit rate was also tested with different threshold window sizes, and the size of the threshold window did not change the significance of the hit rate but only impacted the absolute value of the hit rate.

Table 1.1

Hit Rate of Event Boundaries between Narrational and Narrative Features

	Threshold Window Size (Unit: second)				
	2	**5**	**8**	**10**	**15**
Audial HR	0.1 ***	0.25 ***	0.325 ***	0.325 ***	0.375 ***
Visual HR	0.175 ***	0.275 ***	0.35 ***	0.35 ***	0.475 ***

*** $p < .001$, ** $p < .05$, * $p < .05$. HR = hit rate.

H2. We found support for H2 (**Figure 1.1b**) as the optimal number of events in low-level brain regions (the auditory cortex: 328, the primary visual cortex: 500) was larger than that in high-level brain regions (the medial prefrontal cortex: 84, the precuneus: 197). The optimal number of events in all 17 networks was as follows: 328 in the auditory network[14], 157 in control network A, 112 in control network B, 145 in control network C, 111 in default mode network A, 103 in default mode network B, 247 in default mode network C, 197 in dorsal attention network A, 144 in dorsal attention network B, 191 in the language network, 162 in salience/ventral attention network A, 199 in salience/ventral attention network B, 121 in somatomotor network A, 144 in somatomotor network B, 329 in visual network A, 500 in visual network B[15], and 335 in visual network C.

[14] The auditory network is used as the auditory cortex in this study.
[15] The visual network B is used as the primary visual cortex in this study.

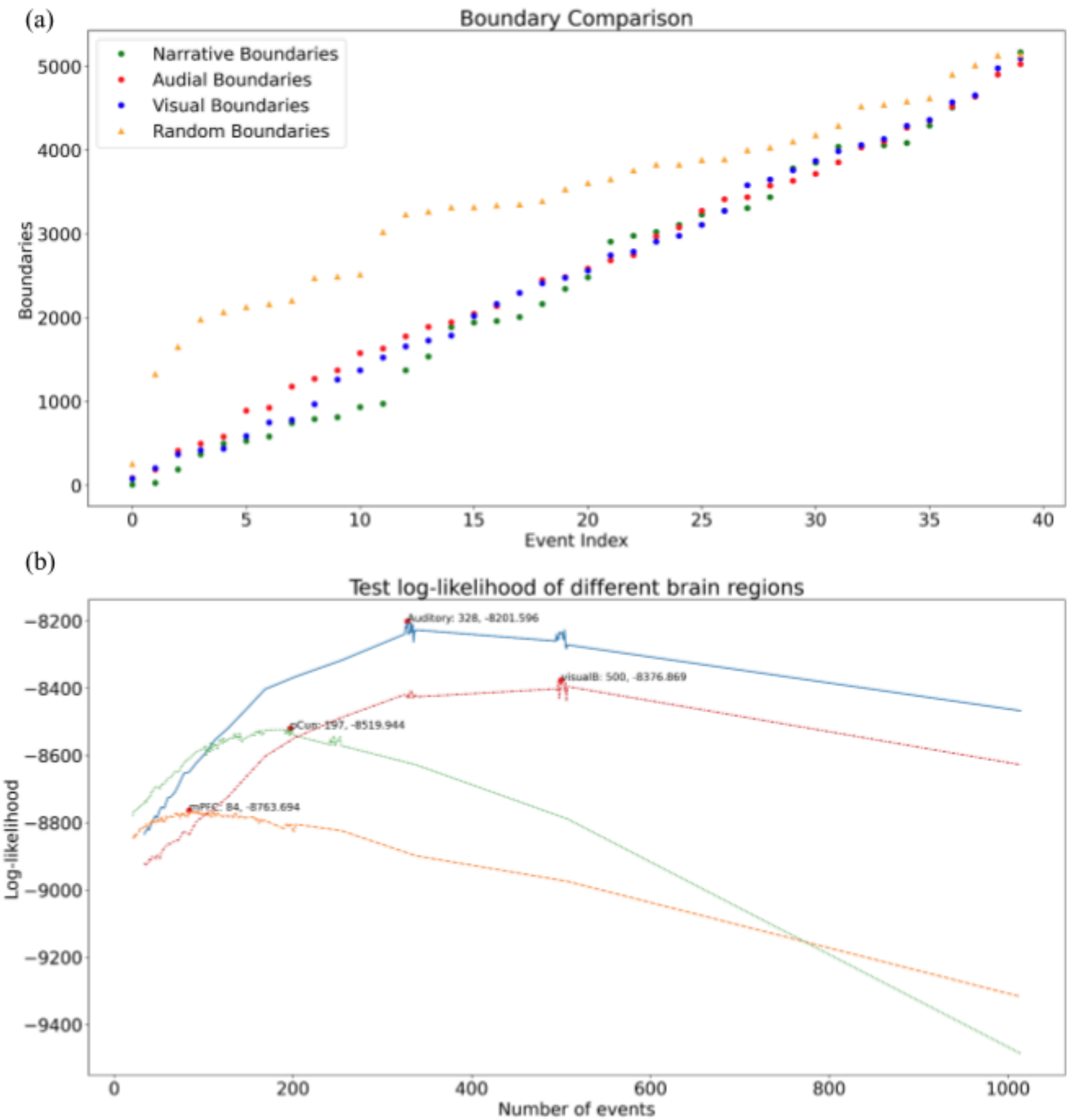


Figure 1.1. (a) Hypothesis 1: Event boundaries in narrative features are correlated with boundaries in narrational features. Red and blue dots are narrational features, green dots are narrative features, and orange triangles are randomly generated narrative features (in the analysis, these triangles are randomly generated 10, 000 times). **(b)** Hypothesis 2: Neural event boundaries are organized in a temporal cortical hierarchy. The optimal number of events for the auditory cortex is 328, with the log-likelihood as -8201.596; for the primary visual cortex (Visual B) is 500, with the log-likelihood as -8376.869; for the medial prefrontal cortex (mPFC) is 84, with the log-likelihood as -8763.694; for the precuneus (pCun) is 197, with the log-likelihood as -8519.944.

H3. We found support for H3 as the hit rates between event boundaries in narrational features and those in low-level brain regions were higher than those in high-level brain regions.

Using a threshold window size of 8, we observed that the hit rates for both audial and visual narrational features were higher in low-level brain regions ($HR_{auditory}$ = 0.832, $p < .001$; HR_{visual} = 0.970, $p < .001$) than in high-level brain regions (HR_{mPFC} = 0.25, $p = .075$; $HR_{precuneus}$ = 0.444, $p < .005$). Similarly, the hit rates for visual narrational features were higher in low-level brain regions ($HR_{auditory}$ = 0.813, $p < .001$; HR_{visual} = 0.976, $p < .001$) than in high-level brain regions (HR_{mPFC} = 0.25, $p = .074$; $HR_{precuneus}$ = 0.538, $p < .001$).

H4. However, our findings did not provide significant support for H4. As displayed in **Table 1.2**, none of the event boundaries in the four brain regions showed a significant association with the boundaries of macro-events in the narrative, as indicated by the insignificant hit rates. We also tested this hypothesis using different threshold window sizes and observed that although the hit rate increased to a relatively large value (e.g., 0.4) as in H3, it remained non-significant.

Table 1.2

Hit Rate of Event Boundaries in Narrative Macro-events with Those in Brain Regions

	Threshold Window Size (Unit: second)		
	10	**20**	**30**
Auditory cortex	0.125	0.225	0.325
Primary visual cortex	0.2	0.325	0.4
Medial PFC	0.1	0.175	0.3
Precuneus	0.1	0.25	0.35

Discussion

The results of Study 1 support the notion that narrational features and narrative features are closely related. Specifically, when there is a shift, change, or break in the narrative features, this boundary is likely to also reflect in the narrational features. This finding suggests that the narration (i.e., how the story is told) plays a crucial role in creating a coherent experience for the audience alongside the narrative (i.e., what the story is about). The time delay captured by the threshold window size indicates that sometimes the narration comes first to prepare the audience for the upcoming narrative, while at other times, it follows the narrative and helps the audience understand what happened. Future studies can further investigate how this specific type of narrative-narration dynamic affects the audience's psychological or neurological states by manipulating different threshold window sizes and types of narrative associated with the shifts.

Our fMRI results also provide evidence to better understand the relationship between narrative and narration. Using a movie that is 10 times longer and contains richer audiovisual information, as well as a different algorithm, we were able to replicate previous findings that low-level brain regions process information faster within a short period, while high-level brain regions process information slower but over a longer period. This finding helps to explain why narrative and narrational boundaries are related but not perfectly matched. When designing narrative and narrational features, content producers must consider that our brain processes these two types of features in dissociable regions and in different ways. It takes time for communication to occur between these regions. Future research can explore how the timing of narrative and narrational features impacts the audience's neural processing and ultimately their experience of the story.

Moreover, we find that the hypothesis on narrational features and low-level brain regions was supported but not on narrative features and high-level brain regions. Our results have shown that low-level brain regions react to narrational features in a way that features are organized; however, high-level brain regions do not respond the same way as the narrative features designed. Previous studies have demonstrated that low-level brain regions process information in a less complicated way than high-level brain regions; therefore, fewer idiosyncrasies are involved. For example, our eyes may pay attention to a colorful painting similarly, but our minds will interpret it entirely differently.

Besides the fact that high-level brain regions have potent capabilities in integrating information and can incorporate out-of-narrative (e.g., personal experience) to facilitate narrative comprehension, other possible reasons explain why our H4 was not supported. First and foremost, the fMRI time series in high-level brain regions is noisier than in low-level brain regions and challenge the accuracy of the algorithm to reach precise results. The movie used in our study is longer than previously published studies and has more data points. We face more complex challenges than in previous studies. Second, the basic idea of HMM is to calculate the probability of a group of information belonging to the same state. The states (i.e., events) in high-level brain regions are longer than in low-level areas. More information brings in more entropy (i.e., uncertainty) and decreases the probability and accuracy. In this vein, the failure to support our H4 is understandable, given the characteristics of the data and the algorithm. Future studies can use other statistical techniques (e.g., Bayesian analysis) to improve the HMM algorithm (Warnick et al., 2018).

In summary, our study 1 results suggest that the narrative and narrational features are related, but not perfectly matched, as they are processed in different brain regions and at different speeds. Our findings highlight the importance of considering both narrative and narrational features when designing content for the audience, as the narration can facilitate the audience's understanding of the narrative. However, our study also shows that high-level brain regions may not respond to the narrative features designed as expected, likely due to the complex information integration involved. Therefore, future studies can explore alternative statistical techniques to improve the accuracy of the algorithm in identifying neural signatures of macro-level mental models. Finally, our findings indicate the need for a more fine-grained approach to understanding the impact of narratives on brain activity in high-level regions. These insights led us to conduct study 2, which investigates how specific narrative features have influenced the micro-level mental model construction within each macro-event.

Study 2 Micro-level Mental Model Construction

Research Questions and Hypotheses

In Study 1, we found that narrational features were strongly associated with low-level brain regions, while narrative features showed more variability, especially in high-level regions. Therefore, in this study, we focus on narrative features and high-level brain regions to gain a more detailed understanding at the micro level.

The results from Study 1 demonstrated that event boundaries detected in high-level brain regions do not perfectly align with narrative features in the movie, which may be attributed to the algorithm's limitations or the brain's tendency to group shorter events (i.e., micro-events) more effectively than longer events (i.e., macro-events). Thus, in Study 2, we aim to investigate the ability of high-level brain regions to perform event segmentation (i.e., mental model construction) at the micro level. Micro-level events refer to the process of event segmentation within macro-events, which involves identifying discrete moments of cognitive processing based on changes in neural activity. Specifically, we test the hypothesis that event boundaries in high-level brain regions are significantly correlated with human-annotated boundaries at the micro level.

H1: *Event boundaries in high-level brain regions are significantly correlated with human-annotated boundaries of narrative events in the movie clips at the micro level.*

Furthermore, not all events are equally important in a movie. Some scenes contain more crucial information (e.g., emotional and moral conflicts) and may evoke different cognitive and emotional states in the audience than other scenes (Hopp et al., 2020; Huskey et al., 2018; Lewis et al., 2014; Tamborini & Weber, 2020; Weber et al., 2012). For example, scenes containing social interactions among characters elicit higher inter-subject correlations between participants' neural activity in the fMRI who watched the same movie (Kauttonen et al., 2015). Similarly, moral information in the movie leads to higher inter-subject correlations among the audience in the insula, cingulate, medial, and lateral prefrontal, superior temporal, and superior parietal cortices (Bacha-Trams et al., 2017).

Inter-subject correlation is a method used in naturalistic data to measure the correlation between BOLD signals from the same brain region of multiple participants. This allows researchers to see how engaged participants are and how much control the stimulus has over their brain activity. Examining inter-subject within each event, either at the macro or micro level, can provide insights to the characteristics of narrative features. For our second hypothesis in study 2, we focus on the effects of morality on the brain at the micro-level. We chose to focus on micro-level moral content because moral judgments are quick and intuitive, and examining shorter segments allows us to avoid noisy information that is morally irrelevant. Our goal is to gain a better understanding of how morally relevant content contributes to the micro-level mental models constructed during movie watching. By exploring the correlation between moral content and inter-subject correlation, we hope to shed light on the specific effects of morality on the brain during movie watching.

Previous studies have established a connection between morally relevant content and inter-subject correlation, which measures the level of engagement and control a stimulus has on the brain. However, the definition of morally relevant content used in these studies was often vague and lacked reproducibility. For example, Bacha-Trams et al. (2017) broadly defined a moral dilemma as "organ donation" in their story. To address this issue, the current study employs the moral foundations theory (Graham et al., 2013) to provide a clear and consistent definition of moral content in movie scenes. The study uses the extended moral foundations dictionary (eMFD) (Hopp et al., 2020) to quantify the moral content of the movie clips. The eMFD was developed based on the moral foundations theory and provides a comprehensive and systematic way to identify moral content in narratives. By applying the eMFD to movie subtitles, this study aims to replicate and extend previous findings (e.g., Bacha-Trams et al., 2017) on the relationship between morally relevant content and neural engagement during movie viewing.

Moral judgments are complex cognitive processes that engage a network of brain regions, including the medial prefrontal cortex, precuneus, and temporal parietal junction (Bzdok et al., 2012; Filley et al., 2020; L. Young & Dungan, 2012). For the purposes of this study, we focus specifically on the medial prefrontal cortex and precuneus, which are known to be key nodes in the default mode network (DMN) (Fransson & Marrelec, 2008). The DMN is a network of brain regions that is most active during rest and self-referential thought and has been implicated in a range of high-level cognitive processes, including social cognition, memory, and imagination (Simony et al., 2016).

H2: *Movie scenes that are higher in morally relevant content have a higher inter-subject correlation in high-level brain regions than non-morally relevant content.*

After replicating the relationship between newly quantified moral content and brain synchronization, this study aims to investigate the relationship between moral content and micro-events. Specifically, if H2 is supported, indicating higher inter-subject correlation in moral scenes than in non-moral scenes, the time-resolved inter-subject correlation (Esfahlani et al., 2020) may provide a proxy of engagement that leads to different results in event segmentation, which is the mental model construction of movie scenes. In Study 1, we used raw BOLD signals for event segmentation, but in this study, we are interested in exploring whether event segmentations detected through the time-resolved inter-subject correlation differ from those detected through raw BOLD signals and quantifying any potential differences.

RQ: *To what extent are micro-level events segmented by the time-resolved inter-subject correlation different from those generated by raw BOLD signals?*

This process of event segmentation is important because it helps us understand how the brain processes information and constructs mental models of events. The time-resolved inter-subject correlation and raw BOLD signals are two commonly used measures of neural activity, but they may generate different micro-level events because they measure neural activity in different ways. The time-resolved inter-subject correlation measures the correlation between the neural activity of different individuals over time, providing information about how brain regions synchronize in response to different stimuli. In contrast, raw BOLD signals measure changes in blood oxygenation levels, which are thought to be related to the level of neural activity in a particular brain region.

It is important to understand whether the two signal sources generate different micro-level events, as this could impact the results of H1 (raw BOLD signals) and H2 (time-resolved inter-subject correlation). By exploring the differences between the two signal sources, we can better understand how to interpret and compare findings from different studies.

Materials and Methods

fMRI Data

Dataset description. This study utilizes fMRI data from two movies, *500 Days of Summer* and *The Grand Budapest Hotel*. The subset data from the Naturalistic Narrative Database (Aliko et al., 2020) was used, with 18 participants watching *500 Days of Summer* after excluding 2 participants due to quality issues noted in Study 1. The functional and anatomical images were obtained using a 1.5 T Siemens MAGNETOM Avanto with a 32-channel head coil, with a TR of 1s. More details about data acquisition and preprocessing can be found in Aliko et al. (2020). The original data can be accessed on OpenNeuro[16]. The data preprocessed with fMRIPrep (Esteban et al., 2019) from (de la Vega et al., 2022) were used in this study. The opening (< 00:00:37, TR < 37) and ending (> 01:26:51, TR > 5211) credits were removed from the movie content data and fMRI data for 500 Days of Summer. The fMRI volumes corresponding to these credits were truncated and not considered for further analysis.

The fMRI data for *The Grand Budapest Hotel* was obtained from a study by Visconti di Oleggio Castello et al. (2020). The full-length movie was divided into six parts of varying durations and watched by 25 participants. The first part, approximately 46 minutes in length, was viewed outside the scanner, while the remaining five parts, ranging from approximately 9 to 13 minutes each, were watched separately in the scanner. All functional and structural volumes were acquired using a 3T Siemens Magnetom Prisma MRI scanner with a 32-channel phased-array head coil, with a TR of 1s. Preprocessing was performed using fMRIPrep, with more details provided in the original study.

For both datasets, we accounted for the delay in hemodynamic response by shifting the fMRI data by 4 TRs, as per Rajapakse et al. (1998).

Regions of interest (ROIs). To identify the medial prefrontal cortex and precuneus, we employed Schaefer et al.'s (2018) 400-parcel 17-network[17] parcellation. For each participant in both the 500 Days of Summer and The Grand Budapest Hotel datasets, we extracted and standardized the time series at the voxel level, and then averaged them into parcel-level data.

[16] https://openneuro.org/datasets/ds002837/versions/2.0.0

[17] 17 networks are the auditory network, the control network (A, B, C), the default mode network (A, B, C), the dorsal attention network (A, B), the language network, the salience/ventral attention network (A, B), the somatomotor network (A, B), and the visual network (A, B, C).

Inter-subject correlation (ISC). After extracting the time series for mPFC and the precuneus, we utilize a leave-one-out approach to calculate the inter-subject correlation separately, by correlating the time series of one participant with the averaged time series of all other participants. This analysis is performed for each movie separately, resulting in unaveraged Pearson's r values at each time point (TR, 1 second) for each participant who watched the same movie (18 participants for *500 Days of Summer* and 25 participants for *The Grand Budapest Hotel*). These results are based on z-scored time series.

Movie Annotations at the Micro Level

Narrative features. We extract moral information from each of the movies using two methods. The first method is the application of the moral foundations dictionary (eMFD) to subtitles. This involves calculating moral scores for each foundation (care/harm, authority/subversion, fairness/cheating, loyalty/betrayal, purity/disgust) based on the presence of moral-foundation-related words in the subtitle of each sentence. Specifically, the dictionary counts moral-foundation-related words for each sentence in the subtitle and returns (1) the score that represents each foundation, (2) the variance across scores for each foundation, and (3) the ratio of moral to non-moral words. It is important to note that this method is applied only to the parts of the movies that have subtitles.

The second method serves as an independent measure and a sanity check and involves human-annotated moral relevance. Two research assistants are trained briefly on moral foundations theory and then asked to rate the extent to which the current scene has moral-relevant content on a scale of 1 to 3. This method is used to confirm that the results obtained from the first method are consistent with the human judgment of moral content in the movies. Previous research suggests that moral content analysis training does not significantly affect the results of such judgments, and moral evaluation is primarily based on intuition (Weber et al., 2018).

Although the eMFD and human-annotated moral relevance measures are different, they both provide useful information about moral content in the movies. The eMFD measure captures moral content based on word usage in subtitles, while the human-annotated measure captures moral content based on human judgment of moral relevance. By using both measures, we can triangulate the moral content in the movies and better understand the relationship between moral content and neural activity. Ultimately, the combination of both measures provides a more robust and nuanced understanding of the moral content in the movies and its relationship to neural activity.

Statistical Analysis

To test H1, which examines the association between event boundaries in the brain and those annotated by human raters at the film's micro level, we use Hidden Markov Modeling (HMM) approach, as in Study 1. For the movie *500 Days of Summer*, we divided the film into 40 macro events using pre-defined calendars. Within each macro event, two research assistants annotated micro events or sub-scenes. We included only the macro events that were long enough (over 5 minutes) to contain two or more sub-scenes for subsequent analyses. We identified four macro events with durations of 398 seconds, 351 seconds, 497 seconds, and 599 seconds, respectively. To analyze the brain signals within the mPFC and precuneus for each macro event, we extracted the corresponding time series. We then used HMMs on the brain data with the number of micro-events annotated by RAs. In cases where the two RAs had different annotation results for the same macro event, we ran multiple HMMs for each RA's result. After obtaining the event boundaries from the HMM, we calculated to what extent the boundaries obtained from

HMM significantly correlated with the boundaries annotated by RAs, using the same method (permutated hit rate, 10,000 permutations in total) described in Study 1.

As described in Study 1, we developed a measurement to compare the closeness between the two types of boundaries. Specifically, for a time point A in the output of the brain HMM, we identify the closest time point B in the hand-annotated boundaries. We then count A as a hit if the difference between A and B is smaller than a given threshold. The threshold value can range from as small as 0 (the most conservative approach) to as large as the length of the shortest event (the most liberal choice in a logical space). The hit rate, which is calculated by dividing the number of hits by the number of events, reflects the strength of the association between event boundaries in the brain HMM and those in hand annotations. The significance of the hit rate is quantified through permutations ($N = 10,000$).

In the case of *The Grand Budapest Hotel*, the original authors segmented the movie into five parts, with each part having a different length ranging from 477 to 782 seconds. Participants watched each part separately while in the scanner. We treat each of these parts as a macro event and two research assistants (RAs) were tasked with annotating micro-events within each macro event. In contrast to the approach taken for *500 Days of Summer*, any discrepancies between the two RAs' annotations were resolved by a third person[18]. We then fit the HMM to the brain data for each part of the movie using the annotated number of micro-events for that part. This results in event boundaries generated by the brain data for each part, which we compare to the boundaries annotated by the RAs using the same method described above.

To test H2, which investigates the relationship between moral-relevant content and inter-subject correlation, we use two sets of generalized linear models. In the first set, we use the

[18] One RA has doubled number of micro-events than the other. After examination, we chose the smaller number of micro-events since these events have more intact meanings rather than pure actions.

variance across moral foundations, as calculated through eMFD, as the independent variable and control for the moral-nonmoral-word ratio and the time point to predict the time-resolved inter-subject correlations. Controlling for the time point is necessary because previous research has shown that inter-subject correlations peak at movie endings when individuals have gathered enough information to reach a shared understanding of the story (Nastase et al., 2019). We only include movie content that contains conversations (i.e., subtitles) in this approach[19].

The second set of generalized linear models has the same dependent variable—the time-resolved inter-subject correlations but a different independent variable—hand-annotated moral-relevant content. We also control for the time in this model. Unlike the first set of models that only include conversations (i.e., subtitles, because eMFD can only apply to words), this approach involves analyzing all parts of the movie.

The above two sets of GLM were applied to each of the two movies.

In order to investigate whether event segmentations from time-resolved inter-subject correlations differ from those generated by averaged BOLD signals (RQ), we utilize the same number of events as used in testing H1 and run HMMs on time-resolved inter-subject correlations. We then calculate the permutated hit rate, as described in Study 1, between boundaries generated by inter-subject correlations and boundaries generated by raw BOLD signals in H1 to quantify any potential differences.

Results

Table 2.1.1

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for 500 Days of Summer (Threshold = 5)

[19] For those parts in the movie that have no conversations have no subtitles/texts to be analyzed.

	# micro-events	Hit Rate (mPFC)	Hit Rate (precuneus)
		Research Assistant #1	
Macro-event 1	2	0	0
Macro-event 2	3	0	0
Macro-event 3	5	0.25**	0
Macro-event 4	5	0	0.25**
		Research Assistant #2	
Macro-event 1	2	0	0
Macro-event 2	4	0	0
Macro-event 3	6	0.2**	0.4***
Macro-event 4	16	0.267***	0.333***

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as five and run 10,000 permutations. The macro-events in *500 Days of Summer* refer to pre-defined editor's cuts longer than 5 minutes. Two different research assistants annotate the number of micro-events.

Table 2.1.2

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for The Grand Budapest Hotel (Threshold = 5)

	# micro-events	Hit Rate (mPFC)	Hit Rate (precuneus)

Macro-event 1	11	0.3***	0.5***
Macro-event 2	4	0	0
Macro-event 3	9	0.25**	0.25**
Macro-event 4	10	0	0.22**
Macro-event 5	8	0	0.143

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as five and run 10,000 permutations. The macro-events in *The Grand Budapest Hotel* refer to pre-divided parts. Event macro-event is the video part watched in the scanner for each run.

H1. The findings from **Tables 2.1.1** and **2.1.2** partially support H1. Specifically, significant results are observed in both *500 Days of Summer* and *The Grand Budapest Hotel* when the number of hand-annotated micro-events is greater than 5. The precuneus shows more significant results than the medial prefrontal cortex. For *500 Days of Summer* (**Tables 2.1.1**), each macro event has a similar time length, but significant relationships between event boundaries in the brain and RA annotations are only observed in macro events with more micro-events within the same time length. Similarly, for *The Grand Budapest Hotel* (**Tables 2.1.2**), each of the five parts has a similar time length, but only those with more micro-events show significant relationships between event boundaries in the brain and RA annotations. Although increasing the hit rate threshold to 10 or 15 yields more significant results (see supplementary materials), the overall pattern remains consistent.

Table 2.2.1

Generalized Linear Model of Effects of Moral Variance (calculated by eMFD) on Time-Resolved Inter-subject Correlations for 500 Days of Summer

	mPFC	Precuneus
	Coefficient (SE)	
Intercept	0.189 (0.061) **	0.257 (0.061) ***
Moral variance	-0.030 (0.031)	0.031 (0.031)
Moral nonmoral ratio	0.009 (0.031)	0.003 (0.031)
Time	-7.577e-05 (2.12e-05) ***	-1e-4 (2.11e-05) ***
No. Observations	1023	1023
Log-Likelihood	-1444.6	-1439.3
Deviance	1009.2	998.65

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.2

Generalized Linear Model of Effects of Moral Variance (calculated by eMFD) on Time-Resolved Inter-subject Correlations for The Grand Budapest Hotel

	mPFC	Precuneus
	Coefficient (SE)	
	Run 1 (Part 1)	
Intercept	0.273 (0.174)	0.030 (0.176)
Moral variance	-0.027 (0.090)	-0.051 (0.091)

Moral nonmoral ratio	0.030 (0.090)	0.047 (0.091)
Time	-0.001 (0.001)	-0.0002 (0.001)
No. Observations	124	124
Log-Likelihood	-174.14	-175.63
Deviance	120.43	123.36
	Run 2 (Part 2)	
Intercept	1.040 (0.367) **	0.257 (0.061) ***
Moral variance	-0.069 (0.128)	0.031 (0.031)
Moral nonmoral ratio	0.244 (0.128)	0.003 (0.031)
Time	-0.003 (0.001) **	-1e-4 (2.11e-05) ***
No. Observations	52	52
Log-Likelihood	-67.426	-65.702
Deviance	40.719	38.105
	Run 3 (Part 3)	
Intercept	-0.274 (0.151)	-0.367 (0.148) *
Moral variance	-0.034 (0.082)	-0.01 (0.08)
Moral nonmoral ratio	0.130 (0.082)	0.184 (0.081) *

Time	0.001 (0.001) **	0.002 (0.001) **
No. Observations	171	171
Log-Likelihood	-239.47	-236.46
Deviance	164.77	159.08
	Run 4 (Part 4)	
Intercept	-0.368 (0.253)	0.237 (0.254)
Moral variance	0.097 (0.100)	-0.072 (0.101)
Moral nonmoral ratio	-0.247 (0.1) *	-0.176 (0.100)
Time	0.001 (0.001)	-0.001 (0.001)
No. Observations	110	110
Log-Likelihood	-152.07	-152.54
Deviance	102.26	103.13
	Run 5 (Part 5)	
Intercept	0.117 (0.216)	0.354 (0.212)
Moral variance	-0.121 (0.099)	-0.148 (0.097)
Moral nonmoral ratio	0.100 (0.099)	-0.034 (0.098)
Time	-0.0003 (0.000)	0.001 (0.000)

No. Observations	104	104
Log-Likelihood	-146.23	-144.56
Deviance	101.35	98.145

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

H2. In our analysis of H2, we did not find significant results to support the hypothesis. We conducted two sets of GLM, with the first using moral variance calculated through eMFD as the predictor and controlling for moral-nonmoral-word-ratio and time. The results, as shown in Table 2.2.1 and Table 2.2.2 for 500 Days of Summer and The Grand Budapest Hotel, respectively, show no significant effect of moral variance on time-resolved inter-subject correlation in the medial prefrontal cortex or the precuneus for either movie. In the second set of GLM, we used hand-annotated moral ratings as the predictor, while controlling for time, and still found no significant results. Further details can be found in the supplementary materials.

RQ. Our findings regarding the differences between event boundaries generated by BOLD signals and time-resolved inter-subject correlations (RQ) are presented in **Table 2.3.1** and **Table 2.3.2**. Our results indicate a similar pattern to what we observed in H1. Specifically, we found that the likelihood of overlap between event boundaries from BOLD signals and time-resolved inter-subject correlations increases with the number of micro-events within a given macro-event. For example, when examining micro-events annotated by research assistant #1, we observed no overlap between the two types of event boundaries (BOLD vs. time-resolved ISC) in either mPFC or precuneus. Even after increasing the threshold from 5 seconds to 15 seconds, we still did not observe a significant overlap. Further details can be found in the supplementary materials.

Table 2.3.1

Hit Rate of Event Boundaries in Averaged BOLD Data and Time-resolved Inter-subject Correlations for 500 Days of Summer (Threshold = 5)

	# micro-events	**Hit Rate (mPFC)**	**Hit Rate (precuneus)**
		Research Assistant #1	
Macro-event 1	2	0	0
Macro-event 2	3	0	0
Macro-event 3	5	0	0
Macro-event 4	5	0	0
		Research Assistant #2	
Macro-event 1	2	0	0
Macro-event 2	4	0	0
Macro-event 3	6	0.2 *	0
Macro-event 4	16	0.2 **	0.13

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as five and run 10,000 permutations. The macro-events in *500 Days of Summer* refer to pre-defined editor's cuts longer than 5 minutes. Two different research assistants annotate the number of micro-events.

Table 2.3.2

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for The Grand Budapest Hotel (Threshold = 5)

	# micro-events	**Hit Rate (mPFC)**	**Hit Rate (precuneus)**
Macro-event 1	11	0.3***	0.4***
Macro-event 2	4	0	0
Macro-event 3	9	0.125	0.375 **
Macro-event 4	10	0.11	0.333 ***
Macro-event 5	8	0	0

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as five and run 10,000 permutations. The macro-events in *The Grand Budapest Hotel* refer to pre-divided parts. Event macro-event is the video part watched in the scanner for each run. Two different research assistants annotate the number of micro-events, and a third person resolves the discrepancies.

Discussion

In this study, we examined the micro-level mental model's two main aspects. Firstly, we investigated whether the event segmentation model is more effective at the micro-level compared to the macro-level in high-level brain regions such as the mPFC and precuneus. Essentially, we sought to determine whether the event boundaries in the brain align with human-annotated boundaries. To achieve this, we developed two event segmentation models: one using BOLD signals averaged across participants and the other using time-resolved inter-subject correlations. We based the two sets on the theoretical argument that they represent distinct narrative characteristics. For instance, ISC is an indicator of engagement, and low ISC coupled with high averaged BOLD signals could suggest that the narrative is not engaging, causing participants' minds to diverge.

Our findings indicate that macro-events with a higher number of micro-events are more likely to display overlapping boundaries in the brain, aligning with the hand-annotated boundaries. This trend holds for both the averaged BOLD signals and time-resolved inter-subject correlations. Moreover, by adjusting the threshold within a reasonable range (without extending too far across events), we observed an increase in the number of significant overlapping boundaries. However, this adjustment did not alter the boundaries that did not have significant overlaps in the first place.

There are three potential reasons why our study did not yield satisfactory overlapping results. Firstly, we compared only two brain regions (mPFC and precuneus) separately with hand-annotated event boundaries. Future studies could compare the results with more theoretically defined brain networks instead of separate brain regions.

Secondly, the higher likelihood of obtaining overlaps when the number of micro-events is larger could be due to statistical issues related to probability theory. For example, given a time series with 1000 data points and two samples, one with one data point and another with ten data points, it is more likely that two data points from the first sample are further apart than 20 data points from the second sample. Future studies could use simulated data to test the event segmentation model's robustness.

Lastly, our significantly different results between event boundaries derived from averaged BOLD signals and time-resolved ISCs confirmed that these signals correspond to different aspects of narrative features. Future studies could use reverse correlation, such as

Ringach and Shapley's method (2004), to quantify different narrative features associated with averaged BOLD signals and ISCs.

The second objective of this study was to investigate the link between key narrative features, specifically moral-relevant content, and brain activity, namely inter-subject correlations in mPFC and precuneus. To achieve this goal, we employed two methods to operationalize moral information: a bag-of-words approach to extract moral information from subtitles and human-annotated moral ratings. However, neither approach yielded the anticipated results that higher levels of moral-relevant content would correspond to higher inter-subject correlations.

There are several potential explanations for the non-significant results. Firstly, with regards to the bag-of-words approach, the eMFD was created using news article data, which has a different corpus and writing style compared to movie subtitles. However, we decided to use this dictionary as it is the most up-to-date and well-established one based on the moral foundations theory. Furthermore, since movies also contain rich visual and auditory information, text-only information from subtitles may only capture a portion of the moral-relevant content.

Although our RA-hand-annotated moral ratings accounted for all visual, auditory, and textual information, we still did not observe a significant relationship between hand-annotated moral information and ISCs. This could be due to two factors: (1) the moral rating procedure and (2) the amount of moral information in the movies. Research assistants were instructed to rate each scene on a scale of 0-3 based on the extent to which it contained moral information. However, this scale may not have been large enough to create a sufficient variance in the ratings. Additionally, it is difficult to clearly define what constitutes moral information, even if all research assistants were familiar with the moral foundations theory. People may judge the same content differently based on their innate moral foundations (Milesi, 2016), meaning that moral ratings from research assistants may not accurately reflect the moral judgments of participants' brains during the movie-watching task in the scanner.

Additionally, we only examined two brain regions, and future studies may need to consider other regions or networks that are involved in moral processing. Nonetheless, our findings provide insight into the construction of micro-level mental models and suggest the need for a more nuanced understanding of narrative processing mechanisms. In Study 3, we will further explore one potential mechanism, namely super-level mental models (i.e., schemas).

Study 3 Super-level Mental Model Construction

Research Questions and Hypotheses

In the previous two studies of this dissertation, we investigated the impact of narrative and narrational features on the construction of mental models at both macro and micro-levels. In this study, we aim to explore how these features contribute to brain activity at the super-level mental models. As defined at the beginning of this dissertation, schemas, and pre-existing mental frameworks, guide our interpretation of stories at the super-level of narrative features by influencing our expectations of events and attention toward relevant information. They strongly influence how we interpret characters' intentions, behaviors, and event sequences. Movies may use ambiguous materials to create suspense, and irrelevant information is often ignored to save cognitive resources. Due to the nature of schemas, which can be either generalized or individualized, it is challenging to create stimuli that evoke the same schema in all participants' brains. To address this, studies such as Aly et al. (2018) have used daily actions (e.g., arriving at a hotel, entering an elevator, boarding an airplane) as a proxy for schemas since they represent knowledge structures that bind information in memory together. These daily actions are highly accessible in the brain and require minimal cognitive effort.

However, it is important to note that there is no guarantee that certain stimuli will evoke the same schema for everyone due to individual and other differences such as cultural and linguistic variations. To address this issue, a naturalistic setting can be used as a possible solution, where stimuli can resemble everyday scenarios, provoking similar schematic thinking among viewers (i.e., engagement, high inter-subject correlation). Although there may still be individual differences in schematically expecting daily actions, it is possible to create stimuli that break schematic thinking for everyone. For instance, a scene that features daily activities, such as a man walking to a restaurant, sitting at a table, and starting to read the menu, preserves a schema. Conversely, a break of the schema would be the man in the scene sleeping on the table or the scene abruptly getting cut off. Therefore, this study takes a comparative approach to uncover the neural underpinnings of schematic thinking by utilizing two naturalistic stimuli: one that preserves schematic thinking and the other that breaks it.

This study utilizes data from Aly et al. (2018), where 30 participants were shown three 90-second clips from the movie The Grand Budapest Hotel: Intact-A, Scrambled-B[20], and Scrambled-C. The aim of this study is to investigate the brain's experience of schema and broken schema, as well as the short-term memory of existing schema and newly formed schema. The first and last watch of clips A (Intact) and C (Scrambled) are used for analysis. The first watch of Intact represents the experience of schema, while the first watch of Scrambled represents broken schema. The last watch of Intact represents the short-term memory of the existing schema, and the last watch of Scrambled represents the short-term memory of the newly formed schema. Further details about the dataset and the choice of clips A and C are provided in the *Material and Methods* section.

The analysis in this study focuses on two brain regions: the precuneus and the primary visual cortex (i.e., the striate cortex). The precuneus was selected due to its critical role in the posterior medial and default mode networks, which are highly involved in constructing complex mental representations (Ritchey & Cooper, 2020). Although the medial prefrontal cortex (mPFC) would also be a valuable region to investigate, fMRI signals in the mPFC were excessively dropped out in the original study. Therefore, it was not possible to include this region in the current analysis. The primary visual cortex was chosen as a control region for the precuneus

[20] As explained later, this one is not used.

since it does not involve any high-level (e.g., cognitive, mentalizing) functions, allowing for a more direct comparison between the two regions.

During the first viewing, we expect that the high-level brain region, precuneus, will respond differently to the intact and scrambled content, while the low-level brain region, the primary visual cortex, will not show any difference between the two clips. We hypothesize that there will be a higher inter-subject correlation in the precuneus for the intact clip, as it follows schematic thinking, compared to the scrambled clip that breaks schemas. On the other hand, inter-subject correlation in the primary visual cortex is expected to be similar in both clips, and significantly greater than zero in both conditions. Therefore, our first hypothesis is to replicate Aly et al. (2018) findings:

H1: *During the first view of the two clips, (a) inter-subject correlation in the precuneus will be significantly higher in the Intact than in the Scrambled; (b) inter-subject correlation in the primary visual cortex will have no difference between the two clips but is significantly greater than zero in both conditions.*

After six repetitions of viewing, participants in Aly et al.'s (2018) study recalled many details about both clips, with slightly better memory for the Intact than Scrambled. One possible explanation is that the repetition enhanced narrative comprehension, leading to increased intra-subject correlation in the precuneus but not the primary visual cortex. This study aims to replicate Aly et al.'s findings.

H2: *In the precuneus, the intra-subject correlation (i.e., correlation between the first and last view) for the Intact and Scrambled clips will be* ***(a)*** *significantly greater than zero and* ***(b)*** *significantly higher for the Intact than the Scrambled clip. In the primary visual* *cortex, the intra-subject correlation for the Intact and Scrambled clips will be* ***(c)*** *not significantly different from each other and* ***(d)*** *not significantly different from zero.*

Moreover, enhanced comprehension in the Intact clip may lead to increased inter-subject correlation since all participants were primed by the same meaningful content. However, this may be different in Scrambled. As the content itself is meaningless, everyone may have a different interpretation, leading to inconsistent interpretations across participants and decreased inter-subject correlation in the last view compared to the first view of the Scrambled clip.

H3: *Inter-subject correlation in the precuneus* ***(a)*** *increases from the first view to the last view of the Intact clip, but* ***(b)*** *decreases from the first view to the last view of the Scrambled clip; inter-subject correlation in the primary visual cortex has no significant changes between the first and the last view in either* ***(c)*** *the Intact or* ***(d)*** *the Scrambled clip.*

Alternatively, the well-performed recall among participants may be due to the learning of the temporal structure through repetition. This would suggest that repeated exposure to the stimuli allowed participants to become more familiar with the sequence and better able to anticipate what was coming next. In this case, we would expect to see increased inter-subject correlation in both the Intact and the Scrambled for the precuneus and the primary visual cortex.

H4: *Inter-subject correlation in the precuneus increases from the first view to the last view of* ***(a)*** *the Intact and* ***(b)*** *the Scrambled; inter-subject correlation in the primary visual cortex increases from the first view to the last view in both* ***(c)*** *the Intact and* ***(d)*** *the Scrambled.*

Table 3.1

Summary of Hypotheses 1 to 4

	Precuneus	Primary visual cortex
	H1	
First-view	**(a)** Inter-subject correlation (Intact > Scambled)	**(b)** Inter-subject correlation (no difference between Intact and Scambled)
	H2	
First-view Last-view	**(a, b)** Intra-subject correlation (Intact > Scrambled > 0)	**(c, d)** Intra-subject correlation (no difference among Intact, Scambled, and 0)
	H3 (the competing hypothesis of H4)	
Intact	**(a)** Inter-subject correlation (First-view < Last-view)	**(c)** Inter-subject correlation (no difference between First-view and Last-view)
Scambled	**(b)** Inter-subject correlation (First-view > Last-view)	**(d)** Inter-subject correlation (no difference between First-view and Last-view)
	H4 (the competing hypothesis of H3)	
Intact	**(a)** Inter-subject correlation (First-view < Last-view)	**(a)** Inter-subject correlation (First-view < Last-view)

Scambled	**(b)** Inter-subject correlation (First-view < Last-view)	**(b)** Inter-subject correlation (First-view < Last-view)

Furthermore, the extent to which a segment in the Scrambled clip deviates from its place in the original movie (how badly the schema was broken) may impact inter-subject correlation. For example, if a segment that was supposed to be at the end of the clip in the original movie was placed at the beginning of the Scrambled clip, it would be considered a large deviation. Similarly, if a segment's position was not far from its place in the original movie, it would be considered a slight deviation. We expect that larger deviations would weaken inter-subject correlation more than minor deviations because larger deviations cause more significant disruptions in the content. This effect is likely to occur in the precuneus, rather than the primary visual cortex, since the former is more involved in meaningful thinking.

H5: *The greater the deviation of a segment from its original place in the movie, the lower its inter-subject correlation in the precuneus will be.*

On the other hand, a deviated segment in the narrative can elicit surprise (e.g., acute mismatches between expectation and reality, Betzel et al., 2017), which can lead to prediction errors and activate specific brain areas such as the precuneus. The more a segment deviates from its original place in the movie, the more surprising it is. However, this effect may diminish over time as the brain realizes the clip is scrambled (i.e., meaningless) and stops generating expectations. Therefore, we have a competing hypothesis for H5:

H6: (a) *The more a segment deviates from its original place in the movie, the higher the inter-subject correlation in the precuneus.* ***(b)*** *This effect may diminish over time.*

H5 and H6 will be tested on the first-view data since the learning effect from repeated viewing can confound the last-view data.

The main goal of this study is to understand the cognitive processes involved in schematic thinking (i.e., viewer engagement), a super-level mental model that enables individuals to organize and interpret information efficiently. Specifically, we aim to investigate how the brain generates and maintains schemas and how it responds to schema violations. By examining the intra- and inter-subject correlation in brain activity during the viewing of intact and scrambled movie clips, we hope to shed light on how the brain processes and represents temporal structure and narrative coherence.

Materials and Methods

fMRI Data

Dataset description. This study utilizes two short clips from *The Grand Budapest Hotel* that depict everyday scenarios. The dataset was obtained from Aly et al. (2018) and consisted of 30 participants (12 male) who viewed three 90-second clips (labeled A, B, and C below) from the movie while undergoing fMRI scanning with a TR of 1.5 seconds. All participants watched all three clips. Clips B and C were created by dividing a 90-second scene from the film into short segments, maintaining natural breaks as much as possible (such as an editor's cut or the end of a spoken sentence), and then rearranging those segments in random orders. Clip B contains 24 segments lasting 2.4 to 5.5 seconds, while Clip C contains 26 segments lasting 2.7 to 4.6 seconds.

In the original study by Aly et al. (2018), 30 participants (12 male) watched three 90-second clips (referred to as A, B, and C below) from The Grand Budapest Hotel while in an fMRI scanner with a TR of 1.5s. All participants viewed all three clips. Clips B and C were created by dividing a 90-second segment of the movie into short segments and then reorganizing those segments with random orders. Clip B contains 24 segments (2.4–5.5 s), and Clip C contains 26 segments (2.7–4.6 s).

For all 30 participants, clip A was viewed in its original format with six repetitions, while Clips B and C were viewed in either a Scrambled-Fixed format (randomized segments were viewed in the same order for all six repetitions) or a Scrambled-Random format (randomized segments were viewed in a different order for each of the six repetitions). Half of the participants viewed Scrambled-Fixed clip B and Scrambled-Random clip C, while the other half viewed the reverse order. Each clip was viewed six times to investigate learning temporal structure over repetitions.

For this study, we use only the first and the sixth (last) views of clips A and C. Clip A contains intact schemas, and clip C contains broken schemas. We chose clip C instead of B because Scrambled-Fixed C was watched right after clip A, while B was watched after C. Using clip C can minimize potential confoundings from not-first-time viewing. Additionally, clip B was quite different from clips A and C in terms of its characters, music, and pace of action.

In summary, our study uses clips A and C that are long enough to trigger schematic thinking by containing meaningful content (e.g., a man getting into an elevator triggers an expectation that he will push a button) while remaining short enough to avoid excessive reliance on movie-specific context information. This data closely mimics the experience of schematic thinking in everyday life, as it involves actions that are common and familiar to the participants and does not require prior knowledge of the movie. Our final dataset comprises the first and sixth views of clip A and the first and sixth views of clip C (Scrambled-Fixed) from 15 participants.

To ensure reproducibility, we use data that was preprocessed by neuroscout (de la Vega et al., 2022) using the fMRIPrep (Esteban et al., 2019).

Regions of interest (ROIs). In line with the methodology employed in Study 1 and Study 2, Study 3 uses Schaefer's (2018) 400-parcel parcellation to identify the precuneus and primary visual cortex and extract the time series for each participant. To account for the delay in hemodynamic response (Rajapakse et al., 1998), all time series are shifted by 3 TRs (4.5 seconds).

Inter-subject correlation. We calculate the correlation between the time series of the precuneus and the primary visual cortex separately for each participant, using a leave-one-out approach. This involves using the time series from one participant to correlate with the averaged time series across all other participants. As a result, we obtain one correlation value (Pearson's r) for each participant for each of the two 90-second (60 TRs) clips.

Intra-subject correlation. The intra-subject correlation between the first-view ROI (i.e., the precuneus and the primary visual cortex) time series and the last-view ROI time series is calculated separately for the intact and scrambled-fixed clips.

Movie Annotations at the Super Level

To quantify the degree to which a segment in clip C deviates from its original temporal structure, we developed an annotation system that involves three steps. First, we annotate the current order of each segment in clip C. Second, we annotate the correct order (order in the original movie) of each segment in clip C (26 segments in total). Finally, we calculate the deviation between the current order and the correct order for each segment. For instance, if the current order of the second segment is 2, but its correct order is 9, the deviation would be 7. The deviation in clip C (Scrambled) ranges from 0 (no deviation) to 25 (the largest deviation) and reflects the extent to which a schema is broken in the scrambled version of clip C.

Statistical Analysis

To test H1 that the first-view inter-subject correlations differ between the intact clip and the scrambled clip for (a) the precuneus and (b) the primary visual cortex, we obtain the first-view inter-subject correlations for both clips and use a paired t-test to compare the difference. Given the small sample size (N=15), the final p-value is corrected through a permutation test (N=10,000) to obtain a more reliable estimation. This analysis is conducted for the precuneus and the primary visual cortex separately.

To test H2, which aims to compare the difference in intra-subject correlation for the Intact and the Scrambled in the precuneus and the primary visual cortex, we use a paired t-test to compare the intra-subject correlations for the Intact and the Scrambled separately in the precuneus and the primary visual cortex. The p-values are corrected using 10,000 permutations.

A similar approach is used to examine the competing hypotheses H3 and H4, which assess the relationship between the inter-subject correlation in both clips' first and last views. To test H3, we calculate the difference for each pair of first-view and last-view correlations in the Intact clip, and then assess the statistical significance with permutation tests (N=10,000). To test H4, we follow the same procedure for the Scrambled clip. The p-values are corrected for multiple comparisons by 10,000 permutations.

To test the competing hypotheses H5 and H6 on the relationship between schema-deviation (variable: deviation) and inter-subject correlation and control for the effect of time (variable: current order), we analyze inter-subject correlations in a time-resolved manner. This involves calculating one Pearson's r for each time point per participant, and then averaging the correlations across participants to obtain a time series that is 60-TRs long. We then compute the mean inter-subject correlation for each segment based on its start and end TRs.

To test H5 and H6a, we set up General Linear Models with inter-subject correlations as the dependent variable and deviation as the predictor for the Scrambled. To test H6b, we include current order as an additional predictor in the model. All models are tested separately for the precuneus and the primary visual cortex, and p-values are corrected for multiple comparisons using permutation tests (N=10,000).

Results

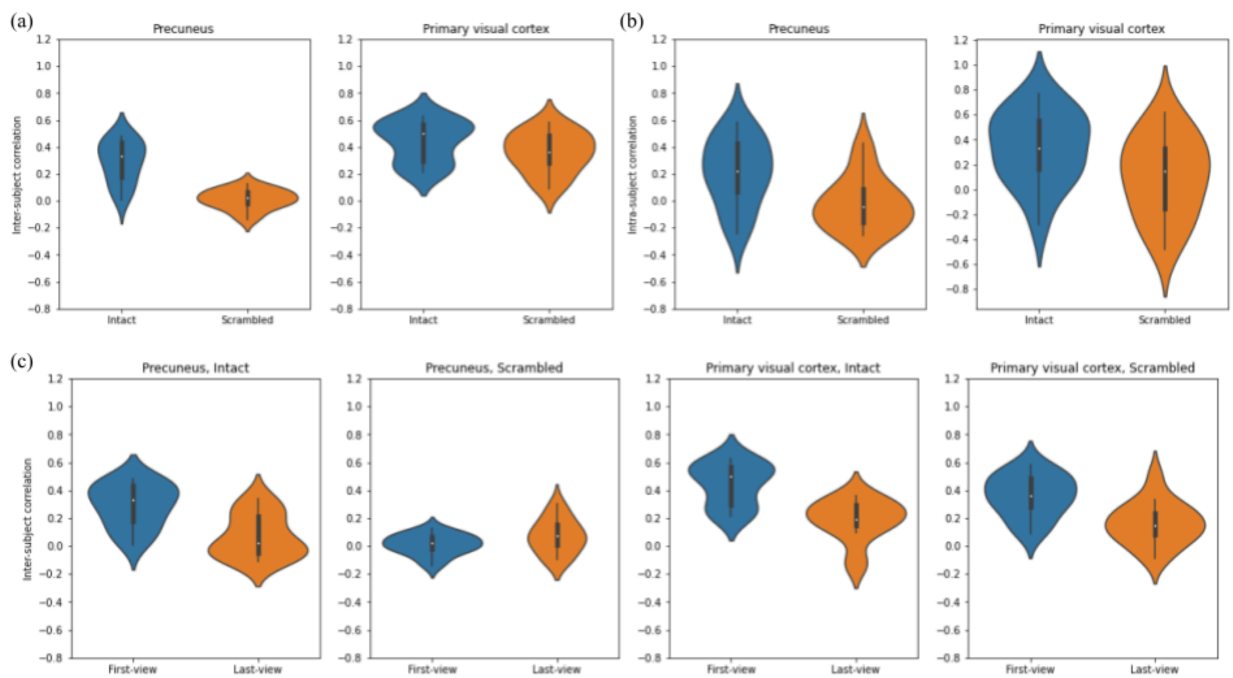


Figure 3.1. (a) Hypothesis 1: Comparing inter-subject correlation in the first view of the two clips. *Left*: In the precuneus, inter-subject correlation for the Intact is significant higher than that for the Scrambled. *Right*: In the primary visual cortex, inter-subject correlation has no significant difference between the Intact and the Scrambled. **(b)** Hypothesis 2: Comparing intra-subject correlation for the two clips. *Left*: In the precuneus, intra-subject correlation is significantly higher for the Intact than the Scrambled. *Right*: In the primary visual cortex, intra-subject correlation is significantly higher for the Intact than the Scrambled. **(c)** Hypotheses 3 & 4: Comparing inter-subject correlation for the first and last views for both clips. *First*: Inter-subject correlation of the first view is significantly higher than that of the last view for the Intact in the precuneus. *Second*: Inter-subject correlation has no difference of the two views for the Scrambled in the precuneus. *Third*: Inter-subject correlation of the first view is significantly higher than that of the last view for the Intact in the primary cortex. *Last*: Inter-subject correlation of the first view is significantly higher than that of the last view for the Scrambled in the primary cortex.

H1. The results of the analysis suggest that H1a is supported, but H1b is not. Specifically, for H1a, in the precuneus, the inter-subject correlation for the Intact ($\text{Inter-SC}_{\text{Intact}}$ = 0.292) is significantly higher than that for the Scrambled-Fixed ($\text{Inter-SC}_{\text{Scrambled}}$ = 0.012, $\text{Inter-SC}_{\text{diff}}$ =

0.280, $t = 6.372$, $p < .0001$) and the inter-subject correlation for the Scrambled is not significantly greater than 0 ($t = 0.634$, $p < .269$). For H1b, opposite to the hypothesis, a significant difference is found in the primary visual cortex ($\text{Inter-SC}_{\text{Intact}} = 0.444$, $\text{Inter-SC}_{\text{Scrambled}} = 0.163$, $\text{Inter-SC}_{\text{diff}} = 0.281$, $t = 5.087$, $p < .0001$) but the inter-subject correlation for the Scrambled-Fixed is significantly greater than 0 ($t = 4.138$, $p < .0001$).

H2. H2a is partially supported, and H2b is not supported. For H2a, in the precuneus, the intra-subject correlation for the Intact ($\text{Intra-SC}_{\text{Intact}} = 0.214$) is significantly higher than that for the Scrambled-Fixed ($\text{Intra-SC}_{\text{Scrambled}} = -0.018$, $\text{Intra-SC}_{\text{diff}} = 0.232$, $t = 2.822$, $p = .005$). In contrast, the intra-subject correlation is not significantly different from 0 ($t = -0.351$, $p = .636$). For H2b, we found that the intra-subject correlation for the Intact ($\text{Intra-SC}_{\text{Intact}} = 0.338$) is significantly higher than that for the Scrambled-Fixed ($\text{Intra-SC}_{\text{Scrambled}} = 0.099$, $\text{Intra-SC}_{\text{diff}} = 0.239$, $t = 2.131$, $p = .022$), which is not significantly different from 0 ($t = 1.186$, $p = .122$).

In testing the competing hypotheses H3 and H4, we found an overall decreased inter-subject correlation from the first to last view in the Intact and the Scrambled for both the precuneus and the primary visual cortex. Specifically, in the precuneus, inter-subject correlation significantly decreased from the first view ($\text{Inter-SC}_{\text{first-view}} = 0.292$) to the last view ($\text{Inter-SC}_{\text{first-view}} = 0.073$) for the Intact ($\text{Inter-SC}_{\text{diff}} = 0.219$, $t = 3.904$, $p < .0001$) but (insignificantly) increased for the Scrambled ($\text{Inter-SC}_{\text{first-view}} = 0.012$, $\text{Inter-SC}_{\text{last-view}} = 0.072$, $\text{Inter-SC}_{\text{diff}} = -0.060$, $t = -1.617$, $p = .943$). In the primary visual cortex, inter-subject correlation significantly decreased from the first view to the last view for both the Intact ($\text{Inter-SC}_{\text{first-view}} = 0.444$, $\text{Inter-SC}_{\text{last-view}} = 0.186$, $\text{Inter-SC}_{\text{diff}} = 0.258$, $t = 4.743$, $p < .0001$) and the Scrambled ($\text{Inter-SC}_{\text{first-view}} = 0.361$, $\text{Inter-SC}_{\text{last-view}} = 0.163$, $\text{Inter-SC}_{\text{diff}} = 0.198$, $t = 3.598$, $p = .001$). Therefore, neither H3 nor H4 is supported.

Our GLM analysis did not provide support for either H5 or H6. Specifically, we did not observe a significant effect of either deviation or current order on inter-subject correlation in the precuneus for the Scrambled stimuli (deviation: $\beta = .01$, SD = .007, $p = .150$, CI = [-0.004, 0.024]; current order: $\beta = -.009$, SD = .005, $p = .07$, CI = [-0.019, 0.001]). We also tested the same model for inter-subject correlation in the primary visual cortex, but did not find any significant effects. However, when we used the averaged time series as the dependent variable, we observed a marginally significant effect of current order on both the precuneus ($\beta = .004$, SD = .001, $p < .001$, CI = [0.002, 0.006]) and the primary visual cortex ($\beta = .005$, SD = .003, $p = .045$, CI = [0.000, 0.011]). Further details can be found in the supplementary material.

When using the averaged time series as the dependent variable, we are measuring the similarity of the signal across subjects at each time point, rather than the similarity of the signal fluctuations across the entire experimental period. In other words, it is a measure of inter-subject correlation at each time point. On the other hand, when using inter-subject correlation as the dependent variable, we are measuring the similarity of the signal fluctuations across the entire experimental period. Therefore, it is possible to observe different results when using these two different dependent variables, as they are measuring different aspects of the signal similarity across subjects.

Discussion

Despite using different approaches to identify the precuneus and the primary visual cortex, our results are still consistent with the original study (Aly et al., 2018) that both intra- and inter-subject correlations in the precuneus are significantly higher for the Intact than that for the Scrambled. Results from the first-view inter-subject correlation support previous findings that the precuneus, as part of posterior medial regions, involves narrative comprehension (Whitney et al., 2009; Wilson et al., 2008; Xu et al., 2005), while the primary visual cortex, as the low-level brain region, does not differentiate between meaningful (the Intact) and meaningless (the Scrambled) content.

Turning to the intra-subject correlation analysis, we observed that, in both the precuneus and the primary visual cortex, the correlation between the first and last views for the Scrambled condition was close to zero. However, interesting differences emerged between the two brain regions when comparing the first- and last-view inter-subject correlations. Specifically, while the last-view inter-subject correlation in the precuneus was slightly higher than the first-view correlation for the Scrambled, this difference was not statistically significant. In contrast, in the primary visual cortex, we observed an insignificant decrease in the inter-subject correlation from the first- to the last-view. Although the lack of statistical significance may be attributed to the small sample size (N=15), these findings suggest that the precuneus may have gradually constructed meaning out of the Scrambled stimuli over the course of six repetitions, leading to a difference in its response between the first and last views. Alternatively, it is possible that the initial processing of the Scrambled stimuli in the primary visual cortex was already quite low, making it difficult to observe significant changes in the inter-subject correlation over time.

In addition, we explored the correlation between the first-view inter-subject correlation and behavioral performance. Surprisingly, we found no significant correlation between the two, neither in the precuneus nor the primary visual cortex. One possible explanation is that the low-level visual processing of the primary visual cortex has little to do with the behavioral task. Participants may only need to see the clips without explicitly attending to them. Therefore, the behavioral performance may not directly reflect the neural response in the primary visual cortex. Another explanation is that the current study only used a short clip, so the behavioral performance may not reflect the entire narrative comprehension. Future studies can employ more extended movie segments to explore this issue further. Overall, our study demonstrates the importance of utilizing both inter- and intra-subject correlation analyses to study neural responses to movie stimuli, which can provide different insights into the neural mechanisms underlying narrative comprehension.

In our GLM analysis, we did not find a significant relationship between the inter-subject correlation and the deviation of a segment (broken schema), despite accounting for the small sample size. There are several possible explanations for this result. First, since participants had no prior experience with The Grand Budapest Hotel or the intact version of the scrambled clip, no shared schemas were established, and thus, no shared deviations occurred. This limitation can be addressed in future studies by incorporating a familiar stimulus. Additionally, even if surprise occurred, participants may have experienced it in different ways, leading to variations in processing surprise. Second, the change in inter-subject correlation may have occurred only at the beginning of a segment, which we have averaged across the entire segment. Future studies could examine the inter-subject correlation at the beginning of each segment to provide a more detailed understanding of the effect of broken schemas on neural activity. Furthermore, our analysis only examined the precuneus, and future research should consider other brain regions, such as the default mode network (DMN), which has established roles in reacting to surprise. Due to our methodological limitations, we could not examine the DMN. Signals from the medial prefrontal cortex (mPFC) have dropped, and we suggest that future studies should examine the DMN to better understand the effects of broken schemas.

Study 3 aimed to investigate how the brokenness of a schema influences inter-subject correlation in naturalistic settings. The findings indicated an insignificant relationship between inter-subject correlation and the deviation of a segment, potentially due to the absence of shared schemas among participants or the variation in surprise processing. Furthermore, the change of inter-subject correlation may only occur at the beginning of a segment, highlighting the need for more fine-designed experiments. Lastly, the study only examined the precuneus and did not explore the role of the DMN in reacting to surprise. Despite the limited findings, the study contributes to the overall theme across all three studies of how the brain processes naturalistic stimuli. It suggests that the brain's response to broken schemas might be a complex and variable process, and future studies should consider the importance of shared schemas, temporal dynamics, and the involvement of other brain regions.

General Discussion

We design three studies in this dissertation to understand macro-, micro-, and super-level mental model constructions during movie viewing and explore the contribution of narrative and narrational features therein. So far, it is the first project that utilizes naturalistic stimuli (i.e., movies) to investigate how the brain processes narrative at three levels. Findings in this dissertation not only help to answer whether narrative and narrational features contribute to mental model construction but also uncover the underlying mechanisms of how they play a role in this hierarchical construction process.

Specifically, Study 1 examined the macro-level mental model construction by investigating how viewers organize and represent the plot and characters of a movie. The results indicated that viewers construct mental models that are consistent with the narrative structure of the movie, which reflects the viewer's ability to integrate plot and character information into a coherent mental representation. These findings suggest that viewers are active participants in constructing mental models that help them comprehend the narrative of the movie.

Study 2 focused on the micro-level mental model construction by examining how viewers represent the visual details of a movie. The results showed that viewers construct mental models that are consistent with the visual features of the movie, which reflects the viewer's ability to process and integrate visual information into a coherent mental representation. These findings suggest that viewers actively engage in constructing mental models that help them process the visual details of the movie.

Study 3 explored the super-level mental model construction by investigating how viewers integrate their prior knowledge with new information to construct mental models during movie-watching. The results revealed that viewers construct mental models that are consistent with their prior knowledge, which reflects the viewer's ability to integrate prior knowledge with new information into a coherent mental representation. These findings suggest that viewers are active in constructing mental models that help them integrate their prior knowledge with new information.

Researchers in communication and media psychology can learn from the neurological component of this dissertation since it offers a biological perspective and methodological innovations to advance our understanding of narration and narrative effects. Next, by connecting narrative and narrational elements with brain activity and tackling narrative processing in three layers rather than one, this dissertation can also be helpful to the neuroscience community. The operationalization of narrative features and links between features or combinations of features and brain activity can be an exemplar for neuroscientists who are less familiar with media studies. Last, the three-level mental model creation can provide filmmakers with insights into how moviegoers typically perceive movies, supported by brain data. When combined with the real-world experience of filmmakers, findings from this dissertation can be of use to facilitate media productions.

Since we use existing datasets in this dissertation, some limitations must be addressed. First, given the availability of the data, those three studies do not use the same movies, so we cannot examine how the three-level comprehension happens simultaneously in the brain. Nevertheless, we try to overcome this drawback by having Studies 1 and 2 share the movie *500 Days of Summer* and Studies 2 and 3 share the movie *The Grand Budapest Hotel*. Second, given the exploratory and original nature of the research questions, the choices of analytical methods remain flexible. In this regard, we propose multiple analytical plans as a method of triangulation for solving some research questions. Last, data from 93 participants are used in all three studies with overlapping subsamples. This number may not be sufficient for generalizability in fMRI research; however, the analytical approach we have proposed here—testing model generalizability across participants, datasets, and events—is still a decent solution to maximize generalizability without thousands of participants (Rosenberg & Finn, 2022).

References

https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Aliko, S., Huang, J., Gheorghiu, F., Meliss, S., & Skipper, J. I. (2020). A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-00680-2

Aly, M., Chen, J., Turk-Browne, N. B., & Hasson, U. (2018). Learning naturalistic temporal structure in the posterior medial network. *Journal of Cognitive Neuroscience*, *30*(9), 1345–1365. https://doi.org/10.1162/jocn_a_01308

Anderson, W. (Director). (2014, March 28). *The Grand Budapest Hotel* [Adventure, Comedy, Crime]. Searchlight Pictures, Indian Paintbrush, Studio Babelsberg.

Andric, M., Goldin-Meadow, S., Small, S. L., & Hasson, U. (2016). Repeated movie viewings produce similar local activity patterns but different network configurations. *NeuroImage*, *142*, 613–627. https://doi.org/10.1016/j.neuroimage.2016.07.061

Armstrong, P. B. (2020). *Stories and the brain: The neuroscience of narrative*. Johns Hopkins University Press.

Bacha-Trams, M., Glerean, E., Dunbar, R., Lahnakoski, J. M., Ryyppö, E., Sams, M., & Jääskeläinen, I. P. (2017). Differential inter-subject correlation of brain activity when kinship is a variable in moral dilemma. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-14323-x

Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, *95*(3), 709-721.e5. https://doi.org/10/gbrrws

Barbas, H., Zikopoulos, B., & Timbie, C. (2011). Sensory pathways and emotional context for action in primate prefrontal cortex. *Biological Psychiatry*, *69*(12), 1133–1139. https://doi.org/10.1016/j.biopsych.2010.08.008

Betzel, R. F., Satterthwaite, T. D., Gold, J. I., & Bassett, D. S. (2017). Positive affect, surprise, and fatigue are correlates of network flexibility. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-00425-z

Brennan, J. R., & Hale, J. T. (2019). Hierarchical structure guides rapid linguistic predictions during naturalistic listening. *PLOS ONE*, *14*(1), e0207741. https://doi.org/10.1371/journal.pone.0207741

Brennan, J. R., Stabler, E. P., Van Wagenen, S. E., Luh, W.-M., & Hale, J. T. (2016). Abstract linguistic structure correlates with temporal activity during naturalistic comprehension. *Brain and Language*, *157–158*, 81–94. https://doi.org/10.1016/j.bandl.2016.04.008

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1–21. https://doi.org/10.1086/448619

Bzdok, D., Schilbach, L., Vogeley, K., Schneider, K., Laird, A. R., Langner, R., & Eickhoff, S. B. (2012). Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. *Brain Structure and Function*, *217*(4), 783–796. https://doi.org/10.1007/s00429-012-0380-y

Chai, L. R., Mattar, M. G., Blank, I. A., Fedorenko, E., & Bassett, D. S. (2016). Functional network dynamics of the language system. *Cerebral Cortex*, *26*(11), 4148–4159. https://doi.org/10.1093/cercor/bhw238

Cutting, J. E., Brunick, K. L., DeLong, J. E., Iricinschi, C., & Candan, A. (2011). Quicker, faster, darker: Changes in Hollywood film over 75 years. *I-Perception*, *2*(6), 569–576. https://doi.org/10.1068/i0441aap

Cutting, J. E., DeLong, J. E., & Nothelfer, C. E. (2010). Attention and the evolution of Hollywood film. *Psychological Science*, *21*(3), 432–439. https://doi.org/10.1177/0956797610361679

de la Vega, A., Rocca, R., Blair, R. W., Markiewicz, C. J., Mentch, J., Kent, J. D., Herholz, P., Ghosh, S. S., Poldrack, R. A., & Yarkoni, T. (2022). Neuroscout, a unified platform for generalizable and reproducible fMRI research. *ELife*, *11*, e79277. https://doi.org/10.7554/eLife.79277

Esfahlani, F. Z., Byrge, L., Tanner, J., Sporns, O., Kennedy, D. P., & Betzel, R. F. (2022). Edge-centric analysis of time-varying functional brain networks with applications in autism spectrum disorder. *NeuroImage*, *263*, 119591. https://doi.org/10.1016/j.neuroimage.2022.119591

Esfahlani, F. Z., Jo, Y., Faskowitz, J., Byrge, L., Kennedy, D. P., Sporns, O., & Betzel, R. F. (2020). High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proceedings of the National Academy of Sciences*, *117*(45), 28393–28401. https://doi.org/10.1073/pnas.2005531117

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), Article 1. https://doi.org/10.1038/s41592-018-0235-4

Filley, C. M., Kletenik, I., & Churchland, P. S. (2020). Morality and the brain: The right hemisphere and doing right. *Cognitive and Behavioral Neurology*, *33*(4), 304–307. https://doi.org/10/gng68g

Finn, E. S., Glerean, E., Hasson, U., & Vanderwal, T. (2022). Naturalistic imaging: The use of ecologically valid conditions to study brain function. *NeuroImage*, *247*, 118776. https://doi.org/10.1016/j.neuroimage.2021.118776

Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *NeuroImage*, *42*(3), 1178–1184. https://doi.org/10.1016/j.neuroimage.2008.05.059

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, *94*(4), 427–438. https://doi.org/10.1037/0033-295X.94.4.427

Geerligs, L., Gözükara, D., Oetringer, D., Campbell, K. L., van Gerven, M., & Güçlü, U. (2022). A partially nested cortical hierarchy of neural states underlies event segmentation in the human brain. *ELife*, *11*, e77430. https://doi.org/10.7554/eLife.77430

Geerligs, L., van Gerven, M., & Güçlü, U. (2021). Detecting neural state transitions underlying event segmentation. *NeuroImage*, *236*, 118085. https://doi.org/10/gkznbh

Gergen, K. J., & Gergen, M. M. (1988). Narrative and the self as relationship. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 21, pp. 17–56). Academic Press. https://doi.org/10.1016/S0065-2601(08)60223-3

Gernsbacher, M. A., Robertson, R. R. W., Palladino, P., & Werner, N. K. (2004). Managing mental representations during narrative comprehension. *Discourse Processes*, *37*(2), 145–164. https://doi.org/10.1207/s15326950dp3702_4

Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. Routledge. https://doi.org/10.4324/9780429500633

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, *53*, 104–114.

https://doi.org/10.1016/j.neuropsychologia.2013.11.010

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. In P. Devine & A. Plant (Eds.), *Advances in Experimental Social Psychology* (Vol. 47, pp. 55–130). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Grosso, A., Cambiaghi, M., Concina, G., Sacco, T., & Sacchetti, B. (2015). Auditory cortex involvement in emotional learning and memory. *Neuroscience*, *299*, 45–55. https://doi.org/10.1016/j.neuroscience.2015.04.068

Hasson, U., Landesman, O., Knappmeyer, B., Vallines, I., Rubin, N., & Heeger, D. J. (2008). Neurocinematics: The neuroscience of film. *Projections*, *2*(1), 1–26.

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

Herman, D. (2004). *Story logic: Problems and possibilities of narrative*. U of Nebraska Press.

Hopp, F. R., Fisher, J. T., Cornell, D., Huskey, R., & Weber, R. (2020). The extended moral foundations dictionary (eMFD): Development and applications of a crowd-sourced approach to extracting moral intuitions from text. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01433-0

Huskey, R., Bowman, N., Eden, A., Grizzard, M., Hahn, L., Lewis, R., Matthews, N., Tamborini, R., Walther, J. B., & Weber, R. (2018). Things we know about media and morality. *Nature Human Behaviour*, *2*(5), 315. https://doi.org/10.1038/s41562-018-0349-9

Jääskeläinen, I. P., Sams, M., Glerean, E., & Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. *NeuroImage*, *224*, 117445.

https://doi.org/10.1016/j.neuroimage.2020.117445

Jang, C., Knight, E. Q., Pae, C., Park, B., Yoon, S.-A., & Park, H.-J. (2017). Individuality manifests in the dynamic reconfiguration of large-scale brain networks during movie viewing. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/srep41414

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge University Press.

Jones, D. (Director). (2011, April 1). *Source Code* [Action, Drama, Mystery]. Summit Entertainment, Vendome Pictures, StudioCanal.

Kauttonen, J., Hlushchuk, Y., & Tikka, P. (2015). Optimizing methods for linking cinematic features to fMRI data. *NeuroImage*, *110*, 136–148. https://doi.org/10.1016/j.neuroimage.2015.01.063

Lahnakoski, J. M., Glerean, E., Jääskeläinen, I. P., Hyönä, J., Hari, R., Sams, M., & Nummenmaa, L. (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage*, *100*, 316–324. https://doi.org/10.1016/j.neuroimage.2014.06.022

Lewis, R. J., Tamborini, R., & Weber, R. (2014). Testing a dual-process model of media enjoyment and appreciation: Media enjoyment and appreciation. *Journal of Communication*, *64*(3), 397–416. https://doi.org/10.1111/jcom.12101

Liu, Z.-Q., Vázquez-Rodríguez, B., Spreng, R. N., Bernhardt, B. C., Betzel, R. F., & Misic, B. (2022). Time-resolved structure-function coupling in brain networks. *Communications Biology*, *5*(1), Article 1. https://doi.org/10.1038/s42003-022-03466-x

Lu, K.-H., Hung, S.-C., Wen, H., Marussich, L., & Liu, Z. (2016). Influences of high-level features, gaze, and scene transitions on the reliability of BOLD responses to natural movie stimuli. *PLOS ONE*, *11*(8), e0161797. https://doi.org/10.1371/journal.pone.0161797

Mar, R. A., Oatley, K., Djikic, M., & Mullin, J. (2011). Emotion and narrative fiction: Interactive influences before, during, and after reading. *Cognition and Emotion*, *25*(5), 818–833. https://doi.org/10.1080/02699931.2010.515151

Masís-Obando, R., Norman, K. A., & Baldassano, C. (2022). Schema representations in distinct brain networks support narrative memory during encoding and retrieval. *ELife*, *11*, e70445. https://doi.org/10.7554/eLife.70445

Meer, J. N. van der, Breakspear, M., Chang, L. J., Sonkusare, S., & Cocchi, L. (2020). Movie viewing elicits rich and reliable brain state dynamics. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-18717-w

Milesi, P. (2016). Moral foundations and political attitudes: The moderating role of political sophistication: MORAL FOUNDATIONS AND POLITICAL SOPHISTICATION. *International Journal of Psychology*, *51*(4), 252–260. https://doi.org/10.1002/ijop.12158

Murphy, S. T., Frank, L. B., Chatterjee, J. S., & Baezconde-Garbanati, L. (2013). Narrative versus nonnarrative: The role of identification, transportation, and emotion in reducing health disparities. *Journal of Communication*, *63*(1), 116–137. https://doi.org/10.1111/jcom.12007

Nash, C. (Ed.). (1994). *Narrative in culture: The uses of storytelling in the sciences, philosophy and literature*. Routledge. https://doi.org/10.4324/9780203981115

Nastase, S. A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, *14*(6), 667–685. https://doi.org/10.1093/scan/nsz037

Nolan, C. (Director). (2010, July 16). *Inception* [Action, Adventure, Sci-Fi]. Warner Bros., Legendary Entertainment, Syncopy.

Ogawa, A. (2021). Time-varying measures of cerebral network centrality correlate with visual saliency during movie watching. *Brain and Behavior*, *11*(9), e2334. https://doi.org/10.1002/brb3.2334

Parrott, W. G. (2002). The functional utility of negative emotions. In *The wisdom in feeling: Psychological processes in emotional intelligence* (pp. 341–359). The Guilford Press.

Pennebaker, J. W., & Seagal, J. D. (1999). Forming a story: The health benefits of narrative. *Journal of Clinical Psychology*, *55*(10), 1243–1254. https://doi.org/10.1002/(SICI)1097-4679(199910)55:10<1243::AID-JCLP6>3.0.CO;2-N

Rajapakse, J. C., Kruggel, F., Maisog, J. M., & Yves von Cramon, D. (1998). Modeling hemodynamic response for analysis of functional MRI time-series. *Human Brain Mapping*, *6*(4), 283–300. https://doi.org/10.1002/(SICI)1097-0193(1998)6:4<283::AID-HBM7>3.0.CO;2-#

Rasmussen, S. (1999). Culture, personhood and narrative: The problem of norms and agency. *Culture & Psychology*, *5*(4), 399–412. https://doi.org/10.1177/1354067X9954002

Regev, M., Honey, C. J., Simony, E., & Hasson, U. (2013). Selective and invariant neural responses to spoken and written narratives. *Journal of Neuroscience*, *33*(40), 15978–15988. https://doi.org/10.1523/JNEUROSCI.1580-13.2013

Ringach, D., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, *28*(2), 147–166. https://doi.org/10.1207/s15516709cog2802_2

Ritchey, M., & Cooper, R. A. (2020). Deconstructing the posterior medial episodic network. *Trends in Cognitive Sciences*, *24*(6), 451–465. https://doi.org/10.1016/j.tics.2020.03.006

Rooney, B., & Hennessy, E. (2013). Actually in the cinema: A field study comparing real 3d and 2d movie patrons' attention, emotion, and film satisfaction. *Media Psychology*, *16*(4), 441–460. https://doi.org/10.1080/15213269.2013.838905

Rosenberg, M. D., & Finn, E. S. (2022). How to establish robust brain–behavior relationships without thousands of individuals. *Nature Neuroscience*, *25*(7), Article 7. https://doi.org/10.1038/s41593-022-01110-9

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, *28*(9), 3095–3114. https://doi.org/10.1093/cercor/bhx179

Schank, R. C., & Abelson, R. P. (2013). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Psychology Press.

Scott, R. (Director). (2012, June 8). *Prometheus* [Adventure, Mystery, Sci-Fi]. Twentieth Century Fox, Dune Entertainment, Scott Free Productions.

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., & Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, *7*(1), 12141. https://doi.org/10.1038/ncomms12141

Song, H., Park, B., Park, H., & Shim, W. M. (2021). Cognitive and neural state dynamics of narrative comprehension. *Journal of Neuroscience*, *41*(43), 8972–8990. https://doi.org/10.1523/JNEUROSCI.0037-21.2021

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences*, *23*(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

Tamborini, R., & Weber, R. (2020). Advancing the model of intuitive morality and exemplars. In K. Floyd & R. Weber (Eds.), *The Handbook of Communication Science and Biology* (1st ed., pp. 456–469). Routledge. https://doi.org/10.4324/9781351235587-39

Tanner, J. C., Faskowitz, J., Byrge, L., Kennedy, D. P., Sporns, O., & Betzel, R. F. (2022). *Synchronous high-amplitude co-fluctuations of functional brain networks during movie-watching* (p. 2022.06.30.497603). bioRxiv. https://doi.org/10.1101/2022.06.30.497603

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The Landscape model of reading: Inferences and the online construction of memory representation. In *The construction of mental representations during reading* (pp. 71–98). Lawrence Erlbaum Associates Publishers.

Vanderwal, T., Eilbott, J., & Castellanos, F. X. (2019). Movies in the magnet: Naturalistic paradigms in developmental functional neuroimaging. *Developmental Cognitive Neuroscience*, *36*, 100600. https://doi.org/10.1016/j.dcn.2018.10.004

Visconti di Oleggio Castello, M., Chauhan, V., Jiahui, G., & Gobbini, M. I. (2020). An fMRI dataset in response to “The Grand Budapest Hotel”, a socially-rich, naturalistic movie. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-00735-4

Warnick, R., Guindani, M., Erhardt, E., Allen, E., Calhoun, V., & Vannucci, M. (2018). A bayesian approach for estimating dynamic functional network connectivity in fMRI data. *Journal of the American Statistical Association*, *113*(521), 134–151. https://doi.org/10.1080/01621459.2017.1379404

Wass, S. V., Smith, C. G., Clackson, K., Gibb, C., Eitzenberger, J., & Mirza, F. U. (2019). Parents mimic and influence their infant’s autonomic state through dynamic affective state matching. *Current Biology*, *29*(14), 2415-2422.e4.

https://doi.org/10.1016/j.cub.2019.06.016

Webb, M. (Director). (2009, August 7). *500 Days of Summer* [Comedy, Drama, Romance]. Searchlight Pictures, Watermark, Dune Entertainment III.

Weber, R., Mangus, J. M., Huskey, R., Hopp, F. R., Amir, O., Swanson, R., Gordon, A., Khooshabeh, P., Hahn, L., & Tamborini, R. (2018). Extracting Latent Moral Information from Text Narratives: Relevance, Challenges, and Solutions. *Communication Methods and Measures*, *12*(2–3), 119–139. https://doi.org/10.1080/19312458.2018.1447656

Weber, R., Popova, L., & Mangus, J. M. (2012). Universal morality, mediated narratives, and neural synchrony. In *Media and the moral mind* (pp. 50–66). Routledge.

Whitney, C., Huber, W., Klann, J., Weis, S., Krach, S., & Kircher, T. (2009). Neural correlates of narrative shifts during auditory story comprehension. *NeuroImage*, *47*(1), 360–366. https://doi.org/10.1016/j.neuroimage.2009.04.037

Whitney, P., Budd, D., Bramucci, R. S., & Crane, R. S. (1995). On babies, bath water, and schemata: A reconsideration of top-down processes in comprehension. *Discourse Processes*, *20*(2), 135–166. https://doi.org/10.1080/01638539509544935

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, *18*(1), 230–242. https://doi.org/10.1093/cercor/bhm049

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Young, C. B., Raz, G., Everaerd, D., Beckmann, C. F., Tendolkar, I., Hendler, T., Fernández, G., & Hermans, E. J. (2017). Dynamic shifts in large-scale brain network balance as a function of arousal. *Journal of Neuroscience*, *37*(2), 281–290. https://doi.org/10.1523/JNEUROSCI.1759-16.2016

Young, L., & Dungan, J. (2012). Where in the brain is morality? Everywhere and maybe nowhere. *Social Neuroscience*, *7*(1), 1–10. https://doi.org/10.1080/17470919.2011.569146

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, *71*, 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

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., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Supplementary Material

Study 2

Table 2.1.1a

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for 500 Days of Summer (Threshold = 10)

	# Segments	Hit Rate (mPFC)	Hit Rate (precuneus)
		Research Assistant #1	
Event 1	2	0	0
Event 2	3	0	å0.5**
Event 3	5	0.25*	0
Event 4	5	0	0.25*
		Research Assistant #2	
Event 1	2	0	0
Event 2	4	0.333**	0
Event 3	6	0.2*	0.4***
Event 4	16	0.333***	0.667***

Note. *** $p < .001$, ** $p < .005$, * $p < .05$ Here, we use the threshold as 10 and run 10,000 permutations. The macro-events in *500 Days of Summer* refer to pre-defined editor's cuts longer than 5 minutes. Two different research assistants annotate the number of micro-events.

Table 2.1.1b

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for 500 Days of Summer (Threshold = 15)

	# Segments	Hit Rate (mPFC)	Hit Rate (precuneus)
		Research Assistant #1	
Event 1	2	0	0
Event 2	3	0	0.5*
Event 3	5	0.25	0
Event 4	5	0	0.25*
		Research Assistant #2	
Event 1	2	0	0
Event 2	4	0.333	0
Event 3	6	0.2	0.4*
Event 4	16	0.6***	0.8***

Note. *** $p < .001$, ** $p < .005$, * $p < .05$ Here, we use the threshold as 15 and run 10,000 permutations. The macro-events in *500 Days of Summer* refer to pre-defined editor's cuts longer than 5 minutes. Two different research assistants annotate the number of micro-events.

Table 2.1.2a

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for The Grand Budapest Hotel (Threshold = 10)

	# Segments	Hit Rate (mPFC)	Hit Rate (precuneus)
Macro-event 1	11	0.5 ***	0.6 ***
Macro-event 2	4	0	0
Macro-event 3	9	0.375 **	0.5 ***
Macro-event 4	10	0.333 **	0.22 **
Macro-event 5	8	0.429 ***	0.143

*** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as 10 and run 10,000 permutations. The macro-events in *The Grand Budapest Hotel* refer to pre-divided parts. Event macro-event is the video part watched in the scanner for each run.

Table 2.1.2b

Hit Rate of Event Boundaries in Averaged BOLD Data and Human-annotated for The Grand Budapest Hotel (Threshold = 15)

	# Segments	Hit Rate (mPFC)	Hit Rate (precuneus)
Macro-event 1	11	0.6 ***	0.6 ***
Macro-event 2	4	0	0
Macro-event 3	9	0.5 **	0.5 ***
Macro-event 4	10	0.444 ***	0.333 **
Macro-event 5	8	0.429 ***	0.286 *

Note.*** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as 15 and run 10,000 permutations. The macro-events in *The Grand Budapest Hotel* refer to pre-divided parts. Event macro-event is the video part watched in the scanner for each run.

Table 2.2.1a

Generalized Linear Model of Effects of Moral Rating on Time-Resolved Inter-subject Correlations for 500 Days of Summer (RA1)

	mPFC	**Precuneus**
	Coefficient (SE)	
Intercept	-0.049 (0.237)	0.157 (0.237)
Moral ratings	-0.096 (0.126)	-0.108 (0.126)
Time	0.000 (0.000)	0.000 (0.000)
No. Observations	66	66
Log-Likelihood	-93.286	-93.039
Deviance	65.276	64.790

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.1b

Generalized Linear Model of Effects of Moral Rating on Time-Resolved Inter-subject Correlations for 500 Days of Summer (RA2)

	mPFC	**Precuneus**
	Coefficient (SE)	
Intercept	0.063 (0.188)	0.152 (0.187)
Moral rating	0.174 (0.100)	0.183 (0.100)
Time	0.000 (0.000)	0.000 (0.000)
No. Observations	100	100
Log-Likelihood	-140.34	-139.9
Deviance	96.945	96.092

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.1c

Generalized Linear Model of Effects of Moral Rating on Time-Resolved Inter-subject Correlations for The Grand Budapest Hotel

	mPFC	**Precuneus**
	Coefficient (SE)	
	Run 1 (Part 1)	
Intercept	0.098 (0.115)	0.240 (0.113) *
Moral rating	-0.015 (0.047)	0.101 (0.047) *
Time	-0.013 (0.015)	-0.033 (0.014) *
No. Observations	578	578
Log-Likelihood	-819.72	-810.91
Deviance	577.15	559.82
	Run 2 (Part 2)	
Intercept	0.805 (0.621)	0.939 (0.619)
Moral rating	0.057 (0.046)	0.088 (0.046)
Time	-0.058 (0.045)	-0.068 (0.045)
No. Observations	478	478

Log-Likelihood	-676.50	-675.00
Deviance	474.54	471.53
	Run 3 (Part 3)	
Intercept	-0.707 (0.388)	-1.197 (0.387)
Moral rating	-0.035 (0.048)	0.046 (0.048)
Time	0.037 (0.020)	0.062 (0.020)
No. Observations	515	515
Log-Likelihood	-727.75	-725.85
Deviance	509.04	505.28
	Run 4 (Part 4)	
Intercept	-1.934 (0.506) ***	-0.375 (0.507)
Moral rating	-0.027 (0.041)	-0.146 (0.041) ***
Time	0.068 (0.018) ***	0.013 (0.018)
No. Observations	598	598
Log-Likelihood	-841.23	-842.26
Deviance	583.59	585.60
	Run 5 (Part 5)	

Intercept	1.571 (0.697) *	2.691 (0.691) ***
Moral rating	-0.065 (0.036)	0.056 (0.036)
Time	-0.046 (0.020) *	-0.078 (0.020) ***
No. Observations	783	783
Log-Likelihood	-1107.3	-1101.3
Deviance	775.61	763.85

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.2a

Generalized Linear Model of Effects of Moral Variance (calculated by eMFD) on Time-Resolved Inter-subject Correlations for 500 Days of Summer

	mPFC	**Precuneus**
	Coefficient (SE)	
Intercept	2.776e-17 (0.031)	-4.944e-17 (0.031)
Moral variance	-0.029 (0.031)	0.031 (0.031)
No. Observations	1023	1023
Log-Likelihood	-1451.1	-1451.1
Deviance	1022.1	1022.0

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.2b

Generalized Linear Model of Effects of Moral Nonmoral Ratio (calculated by eMFD) on Time-Resolved Inter-subject Correlations for 500 Days of Summer

	mPFC	Precuneus
	Coefficient (SE)	
Intercept	2.776e-17 (0.031)	-5.291e-17 (0.031)
Moral nonmoral ratio	0.015 (0.031)	0.015 (0.031)
No. Observations	1023	1023
Log-Likelihood	-1451.5	-1451.5
Deviance	1022.8	1022.8

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.2c

Generalized Linear Model of Effects of Moral Variance (calculated by eMFD) on Time-Resolved Inter-subject Correlations for The Grand Budapest Hotel

	mPFC	Precuneus
	Coefficient (SE)	
	Run 1 (Part 1)	
Intercept	2.29e-16 (0.091)	4.163e-17 (0.090)
Moral variance	-0.015 (0.091)	-0.051 (0.090)
No. Observations	124	124
Log-Likelihood	-175.91	-175.79
Deviance	123.92	123.68
	Run 2 (Part 2)	
Intercept	-9.714e-17 (0.141)	-6.245e-17 (0.141)
Moral variance	-0.089 (0.141)	-0.1030 (0.141)
No. Observations	52	52
Log-Likelihood	-73.578	-73.507
Deviance	51.588	51.4

	Run 3 (Part 3)	
Intercept	-7.286e-17 (0.077)	1.388e-17 (0.077)
Moral variance	0.006 (0.077)	0.046 (0.077)
No. Observations	171	171
Log-Likelihood	-242.64	-242.45
Deviance	170.99	170.63
	Run 4 (Part 4)	
Intercept	-1.492e-16 (0.096)	2.359e-16 (0.095)
Moral variance	0.045 (0.096)	-0.149 (0.095)
No. Observations	110	110
Log-Likelihood	-154.97	-154.84
Deviance	109.78	107.55
	Run 5 (Part 5)	
Intercept	-7.633e-17 (0.098)	8.674e-17 (0.098)
Moral variance	-0.114 (0.098)	-0.142 (0.098)
No. Observations	104	104
Log-Likelihood	-146.89	-146.50

Deviance	102.64	102

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 2.2.2d

Generalized Linear Model of Effects of Moral Nonmoral Ratio (calculated by eMFD) on Time-Resolved Inter-subject Correlations (without controlling) for The Grand Budapest Hotel

	mPFC	**Precuneus**
	Coefficient (SE)	
	Run 1 (Part 1)	
Intercept	2.22e-16 (0.090)	3.469e-17 (0.090)
Moral nonmoral ratio	0.033 (0.090)	0.0474 (0.090)
No. Observations	124	124
Log-Likelihood	-175.88	-175.81
Deviance	123.86	123.72
	Run 2 (Part 2)	
Intercept	-9.714e-17 (0.137)	-7.633e-17 (0.137)
Moral nonmoral ratio	0.243 (0.137)	0.258 (0.137)
No. Observations	52	52
Log-Likelihood	-72.203	-71.992
Deviance	48.930	48.535

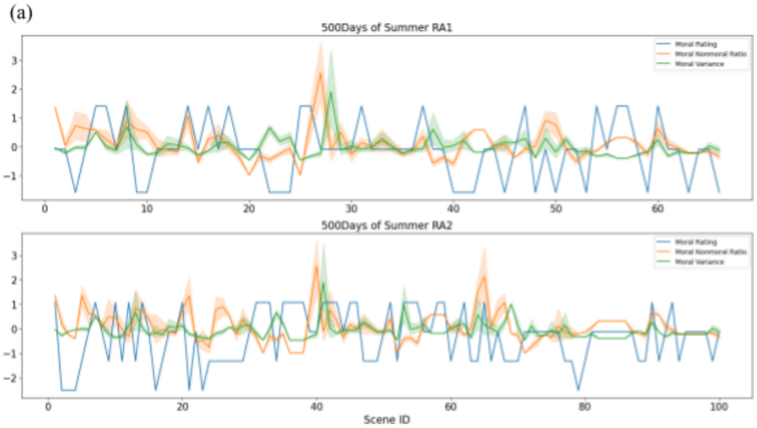
	Run 3 (Part 3)	
Intercept	-7.286e-17 (0.077)	2.082e-17 (0.076)
Moral nonmoral ratio	0.099 (0.077)	0.155 (0.076)
No. Observations	171	171
Log-Likelihood	-241.79	-240.56
Deviance	169.32	166.90
	Run 4 (Part 4)	
Intercept	-1.492e-16 (0.094)	2.359e-16 (0.094)
Moral nonmoral ratio	-0.189 (0.094)	-0.216 (0.094)
No. Observations	110	110
Log-Likelihood	-154.08	-154.45
Deviance	106.07	104.85
	Run 5 (Part 5)	
Intercept	-7.633e-17 (0.099)	9.368e-17 (0.099)
Moral nonmoral ratio	0.089 (0.099)	-0.059 (0.099)
No. Observations	104	104
Log-Likelihood	-147.15	-147.39

Deviance	103.17	103.63

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Figure 2.1

(a)



(b)

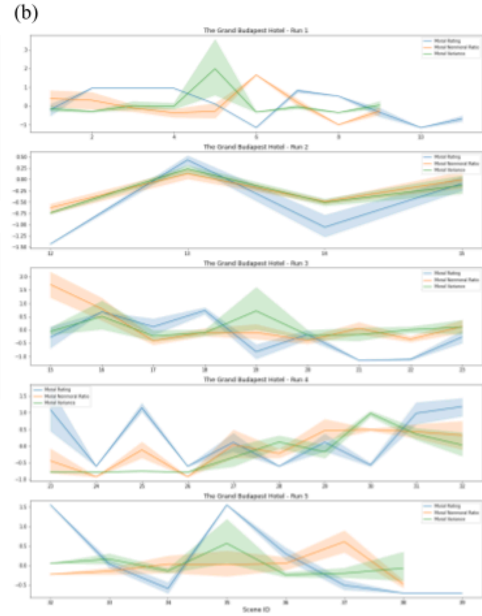


Figure 2.1. **(a)** *500 Days of Summer*. Line plots of moral variance (calculated by eMFD), moral-nonmoral ratio (calculated by eMFD), and hand-annotated moral rating for each scene (defined by each research assistant). Scenes without subtitles have no moral variance and moral-nonmoral ratio. Each subplot is the results from each RA. **(b)** *The Grand Budapest Hotel*. Similar line plots as in **(a)**, but each subplot is the video used in each of the five runs.

Table 2.3.1a

Hit Rate of Event Boundaries in Averaged BOLD Data and Time-resolved Inter-subject Correlations for 500 Days of Summer (Threshold = 10)

	# Micro-events	Hit Rate (mPFC)	Hit Rate (precuneus)
		Research Assistant #1	
Macro-event 1	2	0	0
Macro-event 2	3	0	0
Macro-event 3	5	0	0.25
Macro-event 4	5	0	0
		Research Assistant #2	
Macro-event 1	2	0	0
Macro-event 2	4	0	0
Macro-event 3	6	0.4 *	0
Macro-event 4	16	0.267 **	0.467 ***

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as 10 and run 10,000 permutations. The macro-events in *500 Days of Summer* refer to pre-defined editor's cuts longer than 5 minutes. Two different research assistants annotate the number of micro-events.

Table 2.3.1b

Hit Rate of Event Boundaries in Averaged BOLD Data and Time-resolved Inter-subject Correlations for 500 Days of Summer (Threshold = 15)

	# Micro-events	**Hit Rate (mPFC)**	**Hit Rate (precuneus)**
		Research Assistant #1	
Macro-event 1	2	0	0
Macro-event 2	3	0	0
Macro-event 3	5	0	0.5 **
Macro-event 4	5	0.25	0.25
		Research Assistant #2	
Macro-event 1	2	0	0
Macro-event 2	4	0	0
Macro-event 3	6	0.4 *	0.6 ***
Macro-event 4	16	0.533 ***	0.6 ***

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as 15 and run 10,000 permutations. The macro-events in *500 Days of Summer* refer to pre-defined editor's cuts longer than 5 minutes. Two different research assistants annotate the number of micro-events.

Table 2.3.2a

Hit rate of event boundaries in averaged BOLD data and human-annotated for The Grand Budapest Hotel (Threshold = 10)

	# Micro-events	Hit Rate (mPFC)	Hit Rate (precuneus)
Macro-event 1	11	0.6 ***	0.5 ***
Macro-event 2	4	0	0
Macro-event 3	9	0.375 ***	0.5 **
Macro-event 4	10	0.444 ***	0.444 ***
Macro-event 5	8	0	0.286 **

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as 10 and run 10,000 permutations. The macro-events in *The Grand Budapest Hotel* refer to pre-divided parts. Event macro-event is the video part watched in the scanner for each run. Two different research assistants annotate the number of micro-events, and a third person resolves the discrepancies.

Table 2.3.2b

Hit rate of event boundaries in averaged BOLD data and human-annotated for The Grand Budapest Hotel (Threshold = 15)

	# Micro-events	**Hit Rate (mPFC)**	**Hit Rate (precuneus)**
Macro-event 1	11	0.7 ***	0.5 ***
Macro-event 2	4	0.333 *	0.333
Macro-event 3	9	0.625 ***	0.5 ***
Macro-event 4	10	0.667 ***	0.556 ***
Macro-event 5	8	0	0.429 **

Note. *** $p < .001$, ** $p < .005$, * $p < .05$. Here, we use the threshold as five and run 10,000 permutations. The macro-events in *The Grand Budapest Hotel* refer to pre-divided parts. Event macro-event is the video part watched in the scanner for each run. Two different research assistants annotate the number of micro-events, and a third person resolves the discrepancies.

Study 3

Table 3.1.1

The Effect of Schema Deviation on Inter-subject Correlation in Precuneus

	First-view	**Last-view**
	Coefficient (SE)	
Intercept	0.056 (0.086)	0.044 (0.135)
Deviation	0.010 (0.007)	0.006 (0.011)
Current order	-0.009 (0.005)	-0.001 (0.008)
No. Observations	26	26
Log-likelihood	8.370	-3.463
Deviance	0.800	1.987

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.

Table 3.1.2

The Effect of Schema Deviation on Inter-subject Correlation in Primary Visual Cortex

	First-view	**Last-view**
	Coefficient (SE)	
Intercept	0.253 (0.143)	0.144 (0.157)
Deviation	-0.009 (0.011)	-0.002 (0.013)
Current order	0.012 (0.008)	0.003 (0.009)
No. Observations	26	26
Log-likelihood	-4.816	-7.367
Deviance	2.205	2.683

Note. *** $p < .001$, ** $p < .005$, * $p < .05$.