8-12-2019 2:00 PM

Using movies to assess cognitive and neural functioning in temporal lobe epilepsy

Daniella Ladowski
The University of Western Ontario

Supervisor
Johnsrude, Ingrid
The University of Western Ontario

Graduate Program in Psychology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

Individuals with focal epilepsy whose seizures are poorly managed with medication will often undergo extensive investigations to determine surgical candidacy. These investigations make use of various methodologies to localize normal and pathological brain tissue. Temporal lobe epilepsy (TLE), the most common type of medically refractory epilepsy, can often be detected through structural and functional changes to the affected temporal lobe. On neuropsychological assessment, this dysfunction may be inferred from material-specific memory deficits, with left TLE associated with reduced verbal memory and right TLE associated with reduced visual memory. Although, simple, artificial stimuli may be useful when a clearly lateralizing pattern emerges on testing, other memory deficits may be more subtle or recruit both temporal lobes. Our primary goal with this work was to investigate the utility of a brief, engaging audiovisual film clip to assess temporal-lobe dysfunction in TLE. The first two investigations offer an evaluation of the psychometric properties of a memory test designed to investigate various aspects of memory for the movie. In the first investigation, we used a variety of recall- and recognition-based measures derived from the movie-memory test, whereas the second investigation focused on temporal memory, memory for the temporal context of events in the movie. Both chapters demonstrate the sensitivity of movie-based measures to detect cognitive deficits in TLE. In fact, movie measures appear to be more sensitive than some commonly used standardized tests. The third investigation integrated structural and movie-driven functional neuroimaging measures with performance on the movie-memory test to investigate the combined utility of these methodologies in studying temporal-lobe dysfunction in TLE. Measures of hippocampal volume and connectivity could sensitively distinguish participants with TLE from controls, and abnormal neuroimaging markers could be directly related to cognitive measures to better understand their behavioural consequences. In summary, the current investigations suggest a promising role for movie-based assessment tools in TLE, and motivate their further validation as potential clinical tools to inform surgical planning in TLE.

Keywords

Temporal lobe epilepsy, episodic memory, naturalistic stimuli, neuropsychology

Summary for Lay Audience

Epilepsy is a seizure disorder that can often be treated with medications. However, when medications do not adequately control seizures, brain surgery can be an effective alternative. Surgery consists of removing the part of the brain that is causing seizures, which, in many cases, involves the temporal lobes of the brain. The temporal lobes are important for memory, so although memory may already be affected by epilepsy itself, surgery in this area may cause more substantial memory difficulties. In planning this surgery, different health professionals are asked to identify the part of the brain from which the seizures originate and to consider how a surgery in this area could affect cognitive skills like memory. Neuropsychologists, for example, administer and interpret cognitive tests to make inferences about how well different parts of the brain are functioning. Memory testing typically involves asking the person to learn and remember a series of words or designs. Since everyday memory is more complex, we wanted to investigate whether asking people to remember something more complex and realistic could also be used to assess memory deficits in epilepsy.

We asked people with epilepsy (whose seizures originated in the temporal lobe) and people without any neurological disorder to watch a short, suspenseful movie while they underwent a functional brain scan, and then to complete a memory test for the movie. Their memory for the movie was assessed in different ways, like asking them to state as much as they could remember or asking them to recognize scenes from the movie. We also compared their performance on the memory test with how different parts of the brain were communicating with each other while they watched the movie. We found that the memory test captured memory difficulties in the epilepsy group, the brain scan identified brain differences in the epilepsy group, and together, the memory test and brain scan could be used to clarify how different brain differences manifest as memory difficulties. Future studies can expand on these findings to better understand how tests like these can complement the more traditional tests of memory in presurgical assessments of epilepsy.

Acknowledgments

I am sincerely grateful to my supervisor, Dr. Ingrid Johnsrude, for her wisdom and her warmth. She has nurtured my intellectual curiosity and allowed me to pursue my research interests, and all the way, providing guidance, support, and encouragement. It has been such a privilege to work with her over the past seven years.

I want to thank my other advisory committee members, Dr. Elizabeth Hayden, Dr. Stefan Köhler, and Dr. Ali Khan, for their thoughtful comments and innovative ideas at various stages in the design of this research. Other collaborators, Dr. Mary Pat McAndrews and Dr. Janice Chen, are also owed immense gratitude for their contributions to methodological decisions, and for the body of work to which they have contributed which allowed me to pursue this research.

I am grateful to the clinical and research staff at University Hospital, London Health Sciences Centre. Suzan Brown (research coordinator) has worked tirelessly to recruit participants for this and other EpLink projects, and her dedication to these efforts is extremely admirable. Thanks to Dr. Seyed Mirsattari and Dr. Jorge Burneo (neurologists) for their assistance with recruitment of participants with epilepsy, as well as to Dr. Susan Hayman-Abello and Dr. Brent Hayman-Abello (neuropsychologists) for their contribution of neuropsychological test data.

I am grateful for the support and humour of my colleagues in the CoNCH Lab. I am lucky to have collaborated on research with some, and benefited from the friendship and mentorship of others. We have had a number of excellent undergraduate students in the lab, and this research has certainly benefited from their contributions, especially with respect to data collection and analysis.

Of course, this research would not exist without the numerous participants who volunteered their time. I hope I was able to communicate to them the valuable contribution they provide to the academic community.

Finally, I am humbled by the support (both technical and emotional) of my partner, Logan, and my family, Brenda and Stephanie.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

AED: antiepileptic drug

AI: autobiographical interview

BOLD: blood oxygen level dependent

CALT: Conditional Associative Learning Test

CI: confidence interval

CVLT: California Verbal Learning Test

EEG: electroencephalography

FDR: false discovery rate

fMRI: functional magnetic resonance imaging

HC_{TLE}: healthy control participants – matched to epilepsy sample

HC_{YA}: healthy control participants – young adults

ICV: intracranial volume

iEEG: intracranial electroencephalography

MEG: magnetoencephalography

mPFC: medial prefrontal cortex

MRI: magnetic resonance imaging

MTS: mesial temporal sclerosis

PET: positron emission tomography

RAVLT: Rey Auditory Verbal Learning Test

RCFT: Rey Complex Figure Test

ROI: region of interest

RVDLT: Rey Visual Design Learning Test

rTMS: repetitive transcranial magnetic stimulation

SAM: Survey of Autobiographical Memory

SPECT: single-photon emission computed tomography

TLE: temporal lobe epilepsy (may also refer to the TLE sample)

UWO: University of Western Ontario

WMS-(III or IV): Wechsler Memory Scale-(Third or Fourth Edition)

WAIS-IV: Wechsler Adult Intelligence Scale-Fourth Edition

WASI-II: Wechsler Abbreviated Scale of Intelligence-Second Edition

Chapter 1

1 General introduction

To control epileptic seizures that do not respond to medication, surgical resection of the presumed epileptogenic tissue may be recommended. The surgical work-up involves extensive multidisciplinary investigation to plan a surgical intervention that maximizes the chance of seizure freedom and reduces the likelihood of significant morbidity. Neuropsychological assessment is routinely requested to investigate the individual's cognitive strengths and weaknesses to aid in seizure localization and prediction of cognitive outcomes post-surgically. In some surgical centres, functional magnetic resonance imaging (fMRI) may be used to investigate the neural correlates of cognitive functions (e.g., language and memory) to spare these regions in the resection and/or to predict postsurgical outcomes. For both neuropsychological assessment and fMRI, the cognitive stimuli are typically simple, artificial, and modality specific (verbal or visual) to promote greater specificity of the assessment tools. To complement the information obtained with traditional measures, assessment based on complex, naturalistic, and multimodal stimuli may more closely approximate the demands of everyday cognition and capture aspects of cognition that are missed with simpler stimuli. The potential benefits of enriched stimulation paradigms for surgical planning in refractory epilepsy are discussed.

1.1 Refractory epilepsy

Epilepsy is a chronic neurological condition that affects approximately 1% of individuals in Canada (Tellez-Zenteno, Pondal-Sordo, Matijevic, & Wiebe, 2004). It is characterized by the presence of recurrent seizures, episodes of unprovoked and abnormal neural activity that can result in cognitive, sensory, motor and autonomic disturbances, and occasionally, loss of consciousness. Epileptic seizures may be *focal*, initiated by one or few localized foci in one cerebral hemisphere, or *generalized*, originating simultaneously in both hemispheres and disrupting larger networks of brain activity (Fisher et al., 2017).

Approximately one third of cases of epilepsy are resistant to pharmaceutical treatment (Del Felice et al., 2010; Kwan & Brodie, 2000; Sander, Genton, & Portera-Sanchez, 1993). Epilepsy is formally considered intractable when at least two antiepileptic drug schedules that are appropriately chosen by the physician and adhered to by the patient fail to eliminate seizures (Kwan & Brodie, 2010). When a case of focal epilepsy is deemed intractable, a surgical intervention may be considered.

The cause of epilepsy is an important determinant of intractability and suitability for surgery. Focal epilepsies are more likely to be drug resistant than generalized epilepsies and are more amenable to surgical treatment (Kwan & Brodie, 2000; Mattson, Cramer, & Collins, 1996; Tellez-Zenteno, Dhar, & Wiebe, 2005). The most common type of intractable focal epilepsy is temporal lobe epilepsy (TLE) and the most common cause of TLE is mesial temporal sclerosis (MTS), a loss of neurons and resulting gliosis in the hippocampal formation, including cornu ammonis areas 1 and 3, dentate gyrus, and entorhinal cortex (Briellmann, Kalnins, Berkovic, & Jackson, 2002; Mohanraj & Brodie, 2006). Among individuals with MTS, onset of seizures typically occurs in childhood (Engel, 1996) and may result from prolonged febrile seizures with deleterious effects on the hippocampus (Lewis et al., 2002). As many as 89% of individuals with MTS are not able to achieve seizure freedom despite pharmaceutical intervention (Semah et al., 1998), and approximately 60% of individuals with MTS undergo a temporal resection for seizure reduction (Schuele & Lüders, 2008). Other causes of focal epilepsy that warrant surgical intervention include other lesional (a.k.a. symptomatic) epilepsies (e.g., malformations of cortical development, vascular malformations, tumours, traumatic injury) as well as non-lesional (a.k.a. cryptogenic) epilepsies (when a presumed underlying focal abnormality has not been identified) of temporal and extratemporal origin.

1.2 Preoperative assessment

Two overarching goals guide preoperative assessment in epilepsy. The primary goal is to identify the brain tissue that is thought to generate recurrent seizures and, if removed, is

expected to result in seizure freedom (referred to as the epileptogenic zone or seizure focus). A second goal of the preoperative assessment is to carefully map eloquent cortex, regions of the brain responsible for basic functions like sensation, movement, language, and memory, which, if resected, would lead to substantial postsurgical morbidity. Therefore, the extent of overlap between the epileptogenic zone and eloquent cortex must be determined to weigh the benefit of seizure reduction against the potential cost of functional/neurocognitive sequelae (Rosenow & Lüders, 2001).

Preoperative assessment involves a number of different techniques. The most common tools used for this purpose are summarized in Table 1. The epileptogenic zone is most commonly localized by identifying the approximate location in which whole-brain electroencephalography (EEG) indicates the onset of seizure activity. Video-EEG monitoring, simultaneous recording of electrophysiological data and overt behaviour, captures seizure semiology (behaviour just before, during, and just after electrographic seizures), which can also provide clues as to the location of seizure onset (e.g., déjà-vu feelings early in the seizure are characteristic of temporal-lobe seizures). When epileptiform abnormalities are difficult to detect or localize, alternative electrophysiological recording procedures may be implemented, such as magnetoencephalography (MEG) or intracranial EEG (iEEG), which involves direct recording from the pial surface of the brain. Hypotheses regarding the localization of the epileptogenic zone based on EEG can also be corroborated by searching for structural lesions on magnetic resonance imaging (MRI), focal cerebral metabolic/perfusion disturbances on positron emission tomography (PET) or single-photon emission computed tomography (SPECT), or relative cognitive weaknesses on neuropsychological assessment (from which brain regions of reduced functional integrity can be inferred). Of course, standard neurological practices, including taking a comprehensive history and completing a neurological exam can also provide valuable information about the epilepsy syndrome and localization of the epileptogenic zone (Datta & Loddenkemper, 2011).

Table 1. Commonly used assessment tools for preoperative planning in epilepsy.

Assessment Tool	Description/Target of Assessment	Outcomes of Interest	Alternative/Complementary Tools
EEG	Functional neuroimaging technique measuring the electrical activity in populations of neurons.	EEG recordings are analyzed for epileptiform brain activity. The earliest areas that show epileptiform activity during a clinical seizure may be defined as the *seizure onset zone*. The region that exhibits epileptiform activity between seizures is the *irritative zone* (Datta & Loddenkemper, 2011).	*MEG,* which records the magnetic fields elicited by electrical activity in the brain, may detect spiking activity that EEG cannot (Rodin, Funke, Berg, & Matsuo, 2004). *IEEG* offers higher spatial and temporal resolution than EEG, but because it is invasive, it is reserved for complex localization cases (Yang, Hakimian, & Schwartz, 2014).
Semiology	Signs and symptoms of a seizure including disturbances in sensation, consciousness, motor function, and/or autonomic function (Noachtar, 2001).	Semiology may help to define the *symptomatogenic zone*, which, when activated by seizure activity, causes signs and symptoms characteristic of the seizure. The symptomatogenic zone may overlap or connect to (and hence help to localize) the epileptogenic zone (Rosenow & Lüders, 2001).	*Video-EEG* allows for the simultaneous observation of behavioural signs and recording of neural activity during seizures.

Assessment Tool	Description/Target of Assessment	Outcomes of Interest	Alternative/Complementary Tools
MRI	Structural neuroimaging technique that uses tissue properties to create visual contrast between different anatomical structures.	MRI scans are reviewed for evidence of a brain insult that may reflect the epileptic lesion (e.g., mesial temporal sclerosis, vascular malformation, etc.).	*PET* and *SPECT*, capable of detecting disturbances in cerebral metabolism and perfusion, may help to identify localized abnormalities that are not apparent on MRI.
Wada test (Wada, 1949)	Invasive procedure in which an anaesthetic is administered to one hemisphere via the internal carotid artery, creating a transient lesion. Language and memory tasks are administered to test the residual functioning of the contralateral, "awake" hemisphere.	Anaesthetization of a single hemisphere reveals the lateralization of cognitive abilities such as language and memory. Residual functioning of the "awake" hemisphere informs the extent to which resection of eloquent cortex in the anaesthetized hemisphere would be expected to interfere with cognitive abilities post-surgery.	Less or non-invasive alternatives to the Wada test have been developed including *fMRI* protocols that elicit activation associated with language and memory, and *rTMS*, which creates a targeted, transient lesion, among others (Pelletier, Sauerwein, Lepore, Saint-Amour, & Lassonde, 2007).
Neuropsychological testing	Paper-and-pencil or computerized tasks administered to the patient that target specific cognitive domains.	Cognitive performance can be used to infer areas of functional impairment and functional reserve, to assist with seizure localization and prediction of postoperative cognitive changes (Jones-Gotman et al., 2010).	The patient's subjective reports of cognitive complaints may complement objective test scores.

Note. (i)EEG = (intracranial) electroencephalography; (f)MRI = (functional) magnetic resonance imaging; MEG = magnetoencephalography; PET = positron emission tomography; SPECT = single-photon emission computed tomography; rTMS = repetitive transcranial magnetic stimulation.

Mapping of the eloquent cortex can be accomplished in several ways. Neural activity can be directly or indirectly recorded while the individual performs a specified cognitive task. Functional magnetic resonance imaging (fMRI) is the most common functional neuroimaging methodology used for this purpose, but alternative methods include EEG, MEG, and iEEG. Methods that temporarily excite or inhibit focal neuronal activity (e.g., cortical stimulation, repetitive transcranial magnetic stimulation) can be applied to putative eloquent areas to observe the behaviours that are evoked or suppressed (Loddenkemper & Staudt, 2011). Similarly, Wada testing (Wada, 1949) involves supplying an anaesthetic agent to a single cerebral hemisphere to investigate the preserved cognitive abilities of the contralateral, "awake" hemisphere. Despite its invasive nature, Wada testing is considered the gold standard in lateralizing hemispheric dominance for language, but is also commonly used to investigate relative hemispheric contributions to memory (Baxendale, Thompson, & Duncan, 2008).

Despite the extensive data collected in a routine preoperative assessment, certain presentations can complicate surgical decision making. For example, the location of the epileptogenic zone may be unclear. Assessment may reveal several hypotheses regarding the location of the epileptogenic zone due to discordant information from semiology, structural neuroimaging, functional neuroimaging and neuropsychology (e.g., difficulty identifying the epileptic lesion in a dual/bilateral pathology presentation); or the assessment may be unable to reveal the entirety of the epileptogenic zone due to limitations of imaging resolution (Fountas, 2011; Placantonakis et al., 2010; Voorhies & Cohen-Gadol, 2013; Wellmer et al., 2012). Moreover, the epileptogenic zone may overlap with eloquent cortex. Recall that a primary concern of preoperative assessment is the precise demarcation of eloquent cortex to minimize the risk of serious functional morbidity following resective surgery. In cases of eloquent-cortex involvement, it is important to consider whether functions are already compromised or if surgery poses a significant risk. This point is discussed further below, in reference to the functional adequacy of the affected temporal lobe in TLE and postsurgical memory outcomes.

Localization of eloquent cortex may be particularly complicated if functional reorganization of the cortex has occurred. For example, individuals with TLE, especially those with early-life cerebral insults or longer disease duration, are more likely to exhibit atypical (right or bilateral) hemispheric dominance for language compared to the general population (Branch, Milner, & Rasmussen, 1964; Duchowny et al., 1996; Hamberger & Cole, 2011; Möddel, Lineweaver, Schuele, Reinholz, & Loddenkemper, 2009; Rasmussen & Milner, 1977; Rausch & Walsh, 1984; Springer et al., 1999). In their seminal paper, Rasmussen & Milner (1977) demonstrated the relationship between handedness and hemispheric dominance for language (determined by Wada testing) in individuals with epilepsy associated with early-life left-hemisphere lesions (incurred before age 6) and those without early left-sided lesions. In the group without an early left-sided lesion, 96% of right-handed participants showed left language dominance (4% showed right-sided and 0% showed bilateral organization), and 70% of left- or mixed-handed participants showed left language dominance (15% right, 15% bilateral). In contrast, among the individuals with early left-hemisphere lesions, 81% of right-handed participants showed left language dominance (12% right, 7% bilateral) and a mere 28% of left- or mixed-handed participants showed left language dominance (53% right, 19% bilateral). Thus, it is essential to determine language laterality ahead of surgery to better predict the effect of a resection, especially a temporal-lobe resection, on language and related verbal abilities.

1.3 Neuropsychological assessment in refractory TLE

Unlike other assessment tools that capture neuroanatomical or neurophysiological abnormalities, the neuropsychological assessment measures cognitive abilities, the overt behavioural manifestations of neural changes. The neuropsychological assessment is typically comprehensive, including tests of intellectual functioning, attention, somatosensory and motor functions, visual-spatial skills, language, executive functions, and memory (Jones-Gotman et al., 2010). The individual's overall cognitive profile is inspected. Abilities that are lower than expected (i.e., relative to other abilities, global intellectual functioning, or estimates of premorbid intelligence) that cannot be solely

attributed to secondary influences on cognition (e.g., medication effects, mood, etc.) are used to infer brain regions that may be functionally compromised. Agreement between neuropsychology and other investigations regarding focal abnormalities provides greater confidence regarding the localization of the seizure focus, and suggests that a resection in this region, which already demonstrates evidence of dysfunction, may pose minimal further risk to cognition. Disagreement between neuropsychology and other investigations may trigger further investigation to avoid substantial cognitive deficits postsurgically (Baxendale & Thompson, 2010).

In TLE, episodic memory is the cognitive domain most likely to be impaired presurgically and most at risk of further decline following surgery (typically, an anterior temporal lobectomy) to control seizures (Baxendale, 2008; Hermann, Seidenberg, Haltiner, & Wyler, 1995; Lee, Yip, & Jones-Gotman, 2002; Sabsevitz, Swanson, Morris, Mueller, & Seidenberg, 2001; Spiers, Burgess, Maguire, et al., 2001). Thus, memory is a focus of the neuropsychological assessment. The origins of memory assessment in TLE can be traced back to the influential work of Brenda Milner and her colleagues at the Montreal Neurological Institute. Based on her assessments of individuals with epilepsy, she observed that the left temporal lobe appeared to preferentially process memory for verbal material, whereas the right temporal lobe appeared to preferentially process memory for non-verbal material, and that these material-specific effects were particularly evident following unilateral temporal-lobe resection (Milner, 1970). These early investigations of post-surgical cognitive changes also revealed the potential catastrophic consequences of surgery. The most famous and influential case in this regard is that of H.M., who underwent a bilateral temporal resection (which included portions of both hippocampi) to control seizures, resulting in a global amnesia (Scoville, 1954; Scoville & Milner, 1957). Similarly, other case studies, such as P.B. and F.C., showed that in the context of contralateral temporal-lobe damage, a unilateral temporal lobectomy can also result in global amnesia (Penfield & Mathieson, 1974; Penfield & Milner, 1958). These unacceptable consequences of surgery drive neuropsychologists and the rest of the

health-care team to exhaustively investigate the functional integrity of both temporal lobes.

Following from the seminal contributions of Brenda Milner and others, two basic principles guide the design and selection of memory tests in TLE. One is material-specific effects on memory: the idea that epilepsy originating in the left (or language dominant) temporal lobe is associated with verbal-memory deficits, and epilepsy originating in the right temporal lobe is associated with nonverbal-memory deficits. In TLE, this relationship is most apparent after surgery (Baxendale, 2008; Helmstaedter, Kurthen, Lux, Reuber, & Elger, 2003; Milner, 1970) but has been documented before surgery as well (Gleissner, Helmstaedter, & Elger, 1998; Glosser, Saykin, Deutsch, O'Connor, & Sperling, 1995; Helmstaedter, Pohl, Hufnagel, & Elger, 1991; Hermann, Seidenberg, Schoenfeld, & Davies, 1997; Kim, Yi, Son, & Kim, 2003; Milner, 1972; Sass et al., 1995). The second principle is that both functional adequacy of the affected temporal lobe and functional reserve of the contralateral temporal lobe are predictive of postsurgical morbidity following temporal lobectomy on the affected side (Chelune, 1995). Specifically, intact functioning of the affected temporal lobe is associated with greater risk to memory with surgery, whereas intact functioning of the contralateral temporal lobe is associated with reduced risk to memory. Combining these principles, tests of verbal and visual memory provide information about the functioning of both hemispheres to provide information about seizure localization as well as about likely postsurgical cognitive outcomes.

In selecting tests of verbal and visual memory, Jones-Gotman and an international team working in the neuropsychology of epilepsy (2010) state: "It is best to use tasks that are as purely verbal, or purely nonverbal, as possible to maximize differences between the hemispheres and to increase the probability that the tasks challenge primarily one temporal lobe." (p. 6) Stimuli that favour a single encoding strategy (e.g., words, abstract designs) are generally preferred over dually-encodable stimuli (e.g., audiovisual clips, verbal stimuli that evoke strong imagery, visual stimuli to which verbal labels can be

readily applied). Although simple, unimodal tests have proven their clinical worth since the early days of surgical planning in TLE, relying on such "pure" tests to assess memory function may be restrictive. Unimodal stimuli do not resemble the content of everyday memory. Human memory is complex; simple and arbitrary stimuli are not ecologically valid and may not stimulate cognitive processing the same way that more naturalistic stimuli might.

The material-specificity principle is also not infallible. The link between nonverbal memory and right-hemisphere TLE is more tenuous than the link between verbal memory and left TLE (Baxendale & Thompson, 2010; Bell & Davies, 1998; Glikmann-Johnston et al., 2008; Lee et al., 2002; McAndrews & Cohn, 2012; Saling, 2009). One reason for this may be the heterogeneity among tests that are traditionally considered "nonverbal," with task-specific features dictating which tasks are better at localizing dysfunction and which may recruit bitemporal structures (Glikmann-Johnston et al., 2008). Despite a more consistent relationship between verbal memory and the left temporal lobe, the same issues regarding heterogeneity across verbal-memory tasks and non-exclusive recruitment of left temporal-lobe structures have been debated (Saling, 2009). Furthermore, as stated above, TLE is associated with a higher rate of atypical language representation. Without another investigation (e.g., Wada testing, language fMRI) to lateralize verbal abilities, neuropsychology would be restricted to making inferences about the presumed language-dominant and non-dominant hemispheres instead of left and right.

1.4 MRI in refractory TLE

MRI can be used to collect detailed anatomical images of the brain. The head is positioned in a static magnetic field (typically 1.5 or 3 Tesla) that causes hydrogen atoms in the brain to polarize. An oscillating magnetic field is temporarily applied, causing excitation of the hydrogen atoms. When the hydrogen atoms return to their equilibrium state, they emit radio waves that are detected by sensors. Tissues with different proportions of water and fat (like grey matter, white matter, and cerebrospinal fluid) will emit different signals, creating contrast in the anatomical image.

MRI is a standard investigation in presurgical epilepsy that can be used to identify structural lesions, which typically reflect the underlying epilepsy syndrome (Rosenow & Lüders, 2001). As stated above, the most common structural lesion in TLE is MTS. Evidence of a structural lesion like MTS on presurgical MRI is predictive of significantly better postsurgical seizure relief compared to non-lesional cases of epilepsy (Tellez-Zenteno, Ronquillo, Moien-Afshari, & Wiebe, 2010). The proportion of cases of TLE that are truly non-lesional is hard to know. In about a quarter of TLE cases, a structural lesion cannot be identified on clinical MRI (Carne et al., 2004; Hong, Lee, Kim, Lee, & Chung, 2002). These cases are referred to as "MRI-negative." However, failure to detect a lesion on MRI may, in part, reflect limitations of spatial resolution, as histopathological evidence of MTS has been detected in these presumed non-lesional cases (Kuba et al., 2011; Palacios Bote et al., 2008) and other focal abnormalities can be detected when higher field strengths are used (Mueller et al., 2009; Shah et al., 2019). Reduced effectiveness of surgery in non-lesional cases may be attributed to greater uncertainty regarding the location of epileptogenic tissue.

Volumetric measurement of medial temporal-lobe structures derived from MRI provides a sensitive marker of anatomical changes in TLE. The most precise measurements of hippocampal volume based on MRI can be derived from manual tracing; however, automated segmentation software, like FreeSurfer (http://surfer.nmr.mgh.harvard.edu/), shows good concordance with manual tracing in healthy brains (Cherbuin, Anstey, Réglade-Meslin, & Sachdev, 2009; Morey et al., 2009; Wenger et al., 2014), and is sensitive to medial temporal-lobe atrophy in TLE (Pardoe, Pell, Abbott, & Jackson, 2009). Studies using MRI volumetry consistently show relatively reduced left hippocampal volume in left TLE and right hippocampal volume in right TLE (Barnett, Park, Pipitone, Chakravarty, & McAndrews, 2015; Berkovic et al., 1991; Bernasconi et al., 2003; Doucet, He, Sperling, Sharan, & Tracy, 2016; Fuerst et al., 2001; Lencz et al., 1992; Mechanic-Hamilton et al., 2009). Furthermore, since hippocampal volume is thought to reflect structural integrity, a number of investigators have studied the relationship between reduced hippocampal volume and memory impairment in TLE.

Indeed, left hippocampal volume in left TLE has been linked to verbal-memory abilities although the link between right hippocampal volume in right TLE and nonverbal-memory abilities is not consistently observed (Alessio et al., 2006; Doucet et al., 2016; Glikmann-Johnston et al., 2008; Lencz et al., 1992; Rausch & Babb, 1993).

1.5 fMRI in refractory TLE

When the activity of focal groups of neurons increases, nearby arterioles dilate to increase the flow of oxygenated blood to the region. Oxygenated and deoxygenated hemoglobin in the blood have distinct magnetic properties, and changes in their relative levels evokes a magnetic signal change that is recorded as the blood oxygen level dependent (BOLD) signal. fMRI uses BOLD signal to indirectly measure task- and event-related neuronal activity as individuals perform cognitive tasks.

In TLE, fMRI has become an increasingly routine part of the presurgical work-up. Although fMRI has a number of potential clinical applications in TLE, it is primarily used to lateralize language abilities. For example, the examinee is typically asked to perform expressive and receptive language tasks, such as verbal fluency (e.g., generating words that start with a given letter), responsive naming (generating a word given a verbal description of its meaning), and semantic decision tasks (deciding whether a given word meets certain semantic criteria). The resulting BOLD signal is contrasted in the two hemispheres. When activation is significantly greater in a single hemisphere, this hemisphere is thought to be dominant for language (McAndrews & Cohn, 2012). Protocols like these have shown good concordance with gold-standard techniques for language lateralization (i.e., Wada testing) with the advantage of being non-invasive (Janecek et al., 2013; Szaflarski et al., 2008).

Clinical application of fMRI to localize memory function and evaluate the functional integrity of the medial temporal lobes is promising but remains limited. Some studies have demonstrated concordance between memory fMRI and memory Wada testing in TLE with respect to lateralization of function (Abou-Khalil, 2005; Detre et al., 1998;

Golby et al., 2002), as well as the potential utility of memory fMRI for predicting postoperative memory decline (Binder et al., 2008; Powell et al., 2008; Rabin et al., 2004). The stimuli used in these protocols are comparable to the verbal and visual stimuli used in paper-and-pencil neuropsychological tests. However, fMRI evidence of hemispheric specialization in memory functioning does not appear to neatly follow the material-specific effects that have been observed behaviourally following unilateral temporal-lobe lesions. Kennepohl and colleagues (2007) conducted an fMRI study of encoding and retrieval of verbal (pseudowords, abstract real words) and nonverbal (abstract designs, drawings of objects) stimuli in a healthy sample. They discovered material-specific effects on activation that were independent of hemispheric asymmetries, and hemispheric asymmetries that were not driven by material-specific effects. Furthermore, Binder (2012) argues that targeting specific memory functions with simple stimuli is suitable for addressing questions regarding postoperative memory decline, but may be ill-suited to lateralizing the seizure focus. A task that evokes symmetrical bitemporal activation in healthy samples may be better suited to elucidating asymmetries that implicate one hemisphere over the other.

1.6 Enriched stimulation & the current investigation

To complement the contributions of simple, unimodal stimuli in neuropsychological and neuroimaging investigations in TLE, we sought to investigate the utility of rich, multimodal stimuli that more closely approximate the demands of the everyday situations that our brains have evolved to process and remember.

There has been growing interest in the use of naturalistic stimuli such as movies to understand normal and abnormal brain function (Hasson, Landesman, et al., 2008; Hasson & Honey, 2012; Maguire, 2012). Previous investigators have used movies to simulate real-world memory (Ben-Yakov & Dudai, 2011; Ben-Yakov, Eshel, & Dudai, 2013; Ben-Yakov, Rubinson, & Dudai, 2014; Furman, Dorfman, Hasson, Davachi, & Dudai, 2007; Furman, Mendelsohn, & Dudai, 2012; Hasson, Furman, Clark, Dudai, & Davachi, 2008; Lositsky et al., 2016) and some have used movies to investigate medial

temporal-lobe dysfunction in TLE (Bonnici, Sidhu, Chadwick, Duncan, & Maguire, 2013; St-Laurent, Moscovitch, Jadd, & McAndrews, 2014; St-Laurent, Moscovitch, & McAndrews, 2016).

Investigations based on movie stimuli have begun to reveal aspects of episodic memory and medial temporal-lobe functioning that are not captured using simpler stimuli. For example, the multimodal nature of movies provides a perceptual richness that is known to modulate hippocampal activity (Addis, Moscovitch, Crawley, & McAndrews, 2004; Robin & Moscovitch, 2017; St-Laurent et al., 2016) and reveal the reduced vividness in episodic recall of individuals with TLE (St-Laurent, Moscovitch, et al., 2014). In addition, movies offer a temporal context that is lacking from traditional testing. Different regions of the brain are tuned to time windows of varying lengths, so temporally extended movies allow us to fully explore the contributions of regions that are most sensitive to longer time windows (i.e., on the order of minutes; Hasson, Yang, Vallines, Heeger, & Rubin, 2008; Montchal, Reagh, & Yassa, 2019). Cinematic movies may include a number of natural event boundaries (Baldassano et al., 2017), changes in contextual states over time that can be used to define "episodes" in episodic memory (Clewett, Dubrow, & Davachi, 2019; Kurby & Zacks, 2008; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Event boundaries influence the temporal cohesion of episodic memory (DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011, 2014; Faber & Gennari, 2015; Heusser, Ezzyat, Shiff, & Davachi, 2018; Zwaan, 1996), are known to modulate hippocampal activity (Baldassano et al., 2017; Ben-Yakov & Dudai, 2011; Ben-Yakov et al., 2013; Ben-Yakov & Henson, 2018; Clewett et al., 2019; Ezzyat & Davachi, 2011), and differentially influence memory in individuals with medial temporal-lobe pathology (Zacks, Speer, Vettel, & Jacoby, 2006; although, to date, there have been no investigations of this kind in TLE). Thus, movies are complex enough to engage cognitive processes that are needed for everyday memory but would be obscured by the simple, static nature of traditional cognitive stimuli.

Movies also provide greater experimental control over tests of everyday memory. For example, autobiographical memories are difficult to verify and differ on a number of factors that are known to influence memory, such as the emotional salience of the memory or the complexity of events (Daselaar et al., 2008; St-Laurent, Moscovitch, et al., 2014). Using movies, not only is the stimulus held constant across viewers, but the stimulus itself imposes a kind of control on the mental states of viewers, guiding what they perceive, think, and feel (Hasson, Landesman, et al., 2008), thereby controlling the encoding process. The brain's cortex exhibits predictable, time-locked, and spatially selective activation over the duration of the movie that is synchronized across viewers, quantified as voxelwise inter-subject correlations (Hasson, Nir, Levy, Fuhrmann, & Malach, 2004). The movie stimulus that we selected for the following studies has been shown to evoke stronger and more extensive inter-subject correlations across widespread regions of the cortex, including prefrontal and other default-mode areas that were not reliably activated with other movies (Hasson, Malach, & Heeger, 2010). Greater control over the viewer's attention also appears to translate into reduced head motion for improved sensitivity in neuroimaging analyses (Centeno et al., 2016; Huijbers, Van Dijk, Boenniger, Stirnberg, & Breteler, 2017).

To summarize, movie stimuli have the potential to evoke activity across much of the cortex, reflecting a wide range of cognitive processes engaged, and may also capture specific structure-function relationships that are currently not explored using traditional measures. In the context of presurgical assessments, a stimulus that evokes widespread activity in reliable ways may be used to (a) detect focal functional abnormalities that can assist with localization of the seizure focus, and (b) investigate the functional integrity of numerous regions at once. In addition, tapping novel aspects of medial temporal-lobe functioning that are complementary to traditional tests may enhance the overall sensitivity of the neuropsychological assessment to detect cognitive dysfunction. Thus, movie stimuli may not be as well suited as traditional unimodal stimuli to assessment of lateralization of lesion in epilepsy, but movie-based tools may provide valuable

information about temporal-lobe function in cases of more subtle deficits or when a clearly lateralized pattern does not emerge on traditional tests.

To pursue these ultimate applications of movie-based tools, we must first demonstrate that they are sensitive to the cognitive and neural abnormalities that occur in epilepsy. In the following studies, we investigated whether movie-based assessment tools, including movie-memory testing and movie-driven fMRI could be used to capture episodic memory deficits and medial temporal-lobe dysfunction in TLE. This is the first, exploratory step in validating such tools for clinical use alongside traditional presurgical investigations. A brief, engaging movie stimulus was shown to participants as they underwent fMRI. They also completed a memory test for the movie, composed of different question types designed to maximize sensitivity to medial temporal-lobe dysfunction. Measures derived from the fMRI protocol and memory test were evaluated based on their concordance with traditional cognitive measures and their sensitivity to TLE. In investigating these properties, we hoped to comment on the potential benefit of inclusion of movie-based measures in the presurgical assessment of individuals with refractory TLE.

The following chapters describe the approach that was devised to address the research aims. Chapter 2 describes the phases of data collection. This chapter is intended to orient the reader to the types of data available to pursue the research questions addressed in later chapters. Data from two or more phases of data collection have been incorporated into Chapters 3 through 5. Chapter 3 describes the preliminary psychometric validation of the movie-memory test in healthy and TLE samples. Since aspects of temporal memory emerged as a valuable target for assessment in their own right, Chapter 4 focuses on the validation of these behavioural outcomes specifically in healthy and epilepsy samples. Chapter 5 describes the integration of movie-driven fMRI and movie memory testing as a potential clinical tool for investigating brain-behaviour relationships in TLE. The thesis concludes with a general discussion (Chapter 6).

Chapter 2

2 Phases of data collection

The overarching goal of the current research program was to investigate the potential utility of movie-based assessment tools in the presurgical work-up of individuals with refractory epilepsy. Exploratory investigation of novel behavioural (movie memory testing) and neuroimaging (movie-driven fMRI) paradigms was undertaken. Data collection was completed in three phases: behavioural piloting in a healthy sample (hereafter, piloting), behavioural and neuroimaging validation in a healthy sample (hereafter, healthy validation), and behavioural and neuroimaging validation in a clinical sample (hereafter, clinical validation). This chapter will describe the study design and basic sample characteristics of each phase. In subsequent chapters that address specific research questions, the data from multiple phases are combined to reflect the staged validation process and to improve statistical power. Pertinent information will, therefore, be repeated and elaborated in later chapters, but this chapter may help in orienting the reader to the various datasets available for analysis. See Table 2 for a summary of study procedures by phase of data collection.

2.1 Piloting

2.1.1 Rationale

In the piloting phase, a prototype of the movie-memory test was administered to a group of neurologically healthy young adults. These results were used to provide preliminary evidence of construct validity and to refine the familiarity- and timeline-judgement items to avoid ceiling and floor effects. To test a larger number of items, participants were randomly assigned one of two versions of the memory test, each with a unique set of items for the familiarity- and timeline-judgement tasks. In Chapter 3 (describing the main movie-memory measures), the piloting results for the relevant sections of the memory test are described. In Chapter 4 (describing the temporal-memory measures), the pilot testing is mentioned in describing the selection of timeline-judgement items.

Table 2. Study procedures.

Study	Neuroimaging	Movie Memory Test	Standardized Neuropsychological Tests	Questionnaires
Behavioural Pilot	N/A	Original Prototype, version A or B	Verbal memory: Logical Memory Visual memory: Rey Complex Figure Test, Conditional Associative Learning Test Other: Digit Span	
Healthy Validation (HC_{YA})	Movie-driven fMRI	Revised Test	Verbal memory: Logical Memory, Rey Auditory Verbal Learning Test Visual memory: Conditional Associative Learning Test Intelligence: Matrix Reasoning	Survey of Autobiographical Memory, Episodic Subscale
Clinical Validation (TLE, HC_{TLE})	Movie-driven fMRI	Revised Test	Verbal memory: Logical Memory (TLE only), California Verbal Learning Test (TLE only), Names Test Visual memory: Rey Complex Figure Test (TLE only), Conditional Associative Learning Test, Rey Visual Design Learning Test, Doors Test Intelligence: Matrix Reasoning, Vocabulary	

Note. HC_{YA} = healthy control participants – young adults; HC_{TLE} = healthy control participants – matched to epilepsy sample; TLE = temporal lobe epilepsy sample.

2.1.2 Participants

Neurologically healthy individuals were recruited from among the undergraduate population at the University of Western Ontario (UWO) using a participant recruitment database. Thirty participants (age $M \pm SD = 18.63 \pm 0.96$; 21 female) were enrolled, all of whom self-identified as native English speakers with no history of hearing impairment or neurological disorder. All participants provided informed consent and received course credit in exchange for participation.

2.1.3 Procedure

After viewing the movie stimulus on laptop, participants completed four types of movie-memory questions: the free- and probed-recall interview, familiarity-judgement task (24 items), timeline-judgement task (20 items), and general comprehension questions (20 questions). The interview was performed aloud, whereas the remaining sections were administered on laptop, using the Psychophysics toolbox (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997) in MATLAB 2014b (Mathworks, Natick, MA, USA). Examples of each question type are shown in Figure 1. With each section, we aimed to assess different aspects of memory for the movie. Since the movie was multimodal and can be encoded using visual and/or verbal strategies, different question types were expected to facilitate visual, verbal, or combined visual and verbal retrieval strategies. Participants also completed a brief battery of standardized neuropsychological measures.

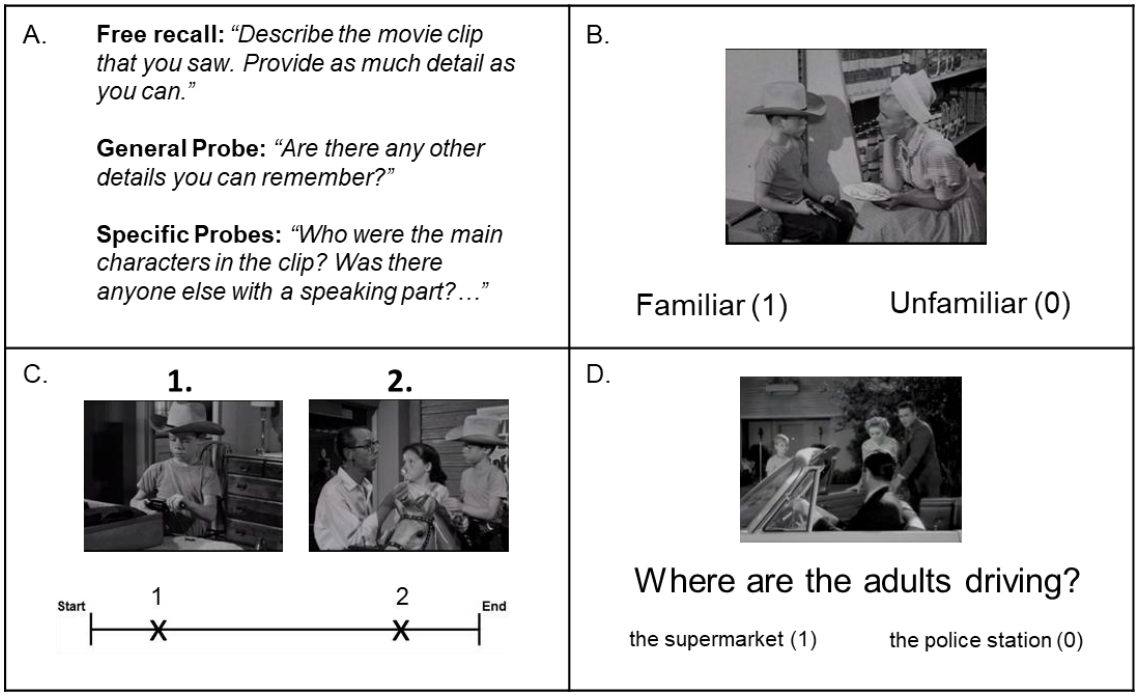


Figure 1. Sample items for each type of memory question. *A* depicts an abridged version of the questions that were asked in each of the recall conditions (free recall, general probe, specific probes). For the complete interview, see Appendix A. *B* depicts a sample item from familiarity judgements, for which the participant was asked to type a 1 of 0 depending on their selection. *C* depicts a sample item from timeline judgements. The timeline initially appears blank, with just the "Start" and "End" anchors. When the participant responds, an "X" appears on the timeline in the spot where the participant judged the item to have occurred, labeled with the number (1 or 2) corresponding to the still frame that was judged. *D* depicts a sample comprehension question, for which the participant was asked to type 1 of 0 depending on their selection. For the complete list of questions, see Appendix D.

2.1.3.1 Free & probed recall

Free and probed recall for the movie followed a structured interview format, adapted from the Autobiographical Interview (AI; Levine, Svoboda, Hay, Winocur, & Moscovitch, 2002). In free recall, participants were instructed to provide as much detail

as they could remember about the movie. In probed recall, participants were prompted to provide any more information that they could recall (general probe) and were then asked a series of questions pertaining to various aspects of the movie, such as the characters, settings, and perceptual details (specific probes). No time limit was imposed. The complete interview is available in Appendix A. To score free and probed recall, a checklist of events and other detail types was devised from the plot of the movie. For each participant, the number of correctly reported details in the checklist were tallied.

2.1.3.2 Familiarity judgements

For familiarity judgements, participants were shown a single still frame and were asked to judge whether it was familiar or unfamiliar to them. They were subsequently asked to rate their confidence on a four-point scale, from 1-very unsure to 4-very confident. Twenty-four items were presented in total, 12 targets and 12 lures. Forty-eight items were piloted across the two versions. An equal number of target items was sampled from each of the two-minute quarters of the movie to ensure adequate sampling from the entire length of the movie. As the eight-minute movie was edited down from a 20-minute television episode, distractor items could be sourced from the unused content of the movie clip, meaning that characters, settings, and scenes are often the same as those presented in the abridged clip. The main outcome measure was the sensitivity of the participant's recognition memory, indexed by d'.

2.1.3.3 Timeline judgements

For timeline judgements, participants were shown two still frames side-by-side as well as a timeline, for which the only anchors provided were the start (extreme left) and end (extreme right) points of the movie. They were instructed to click on the timeline where they recalled each still frame to have taken place in the course of the movie (i.e., two clicks per item). Twenty items, composed of 40 still frames, were presented in total. Forty items were piloted across the two versions. In each version, to ensure the entirety of the movie was sampled, 10 still frames were sampled from each quarter (i.e., 120 s segment) of the movie. Additionally, the still frames for each item were sampled at

different intervals, including 240 s apart, 120 s, 60 s, 30 s, and 15 s (4 items per interval length). For half of the items, the two still frames were ordered as they were in the movie (i.e., earlier frame on the left). Scoring was based on a single outcome measure, position estimation error. For each still frame, the absolute difference between the position on the timeline where the participant clicked and its true position was computed, converted to seconds, and averaged across all still frames.

2.1.3.4 Comprehension questions

General comprehension questions were constructed to assess whether the participant was following the plot of the movie. The questions took a two-option forced choice format and, when appropriate, were accompanied by a still frame from the movie to provide contextual support. Some of the questions could only be answered by recalling specific segments of the movie on the order of seconds, and these segment-specific questions were sampled from the full length of the movie. They were again asked to rate their confidence on a four-point scale for each question. Twenty items were presented in total (see Appendix D for a complete list of questions). The total number of correct responses was tallied.

2.1.3.5 Neuropsychological tests

The Logical Memory subtest of the Wechsler Memory Scale-Fourth Edition (Wechsler, 2009) involves learning and delayed recall of two short stories. The delayed recall measure was used for analysis. The Rey Complex Figure Test (RCFT; Strauss, Sherman, & Spreen, 2006b) involves the copy and delayed recall of a complex abstract figure. Again, the delayed recall measure was used for analysis. The Conditional Associative Learning Test (CALT; Petrides, 1985; St-Laurent, McCormick, et al., 2014) involves learning the arbitrary associations between cards and spatially distributed discs with examiner feedback. The total number of trials administered before the criterion to discontinue was met (12 consecutive correct, or a maximum of 64 trials) was used for analysis. The Digit Span subtest of the Wechsler Adult Intelligence Scale-Fourth Edition

(Wechsler, 2008; including forward, backward, and sequencing conditions) was administered as a measure of auditory attention and working memory.

2.2 Healthy validation

2.2.1 Rationale

In the healthy-validation phase, after revising the movie-memory test based on the results of piloting, the revised version was administered to a sample of neurologically healthy young adults. None of these individuals had participated in the piloting phase. In addition to improving the overall statistical power of analyses involving the clinical-validation samples when appropriate, the healthy-validation data were used to test the inter-rater reliability of scoring guidelines that were constructed for the free- and probed-recall interview. These results were then used to revise the scoring guidelines and to refine the study variables to only include those with adequate inter-rater reliability. Data from this phase have been incorporated into Chapters 3 (movie memory), 4 (temporal memory for the movie), and 5 (neuroimaging and memory).

2.2.2 Participants

Neurologically healthy young adults (referred to as HC_{YA}) were recruited from the UWO and greater London community through flyers. Twenty-four participants (age $M \pm SD$ = 23.17 ± 3.24; years of education $M \pm SD$ = 15.63 ± 2.32; 14 female) were enrolled, all of whom self-identified as native English speakers with no history of hearing impairment, psychiatric illness, or neurological disorder. All participants provided informed consent and received nominal compensation to participate.

2.2.3 Procedure

Participants viewed the movie while undergoing functional magnetic resonance imaging (fMRI). The fMRI scan was a continuous acquisition, similar to the procedure typically followed for resting-state fMRI. A structural scan was also obtained. Upon exiting the scanner, participants completed the movie-memory test (consisting of the free- and

probed-recall interview, 16 familiarity-judgement items, 28 timeline-judgement items, and 20 comprehension questions) as well as a brief battery of standardized neuropsychological tests.

2.2.3.1 Free & probed recall

This section followed the same interview format as described for piloting. However, instead of scoring recall based on a detail checklist, a more sophisticated scoring system was devised. Since the interview format was based on the AI, it was felt that the scoring of the AI could also be adapted for the movie stimulus. In fact, two scoring systems were devised, one that was directly based on the AI and derivatives thereof (St-Laurent, Moscovitch, et al., 2014; used in Chapter 3) and one that was based on a derivative of the AI (St-Laurent, Moscovitch, Tau, & McAndrews, 2011; used in Chapter 4) that could be used to investigate temporal aspects of episodic memory. Two independent raters used a preliminary version of the scoring guidelines to rate each of the 24 recall transcripts. Disagreements across raters were discussed, and the scoring guidelines were revised accordingly. The final scoring guidelines are available in Appendix B (general scoring instructions) and Appendix C (temporal scoring instructions).

2.2.3.2 Familiarity judgements

This section followed the same procedure as described for piloting. A set of 16 items was selected, balancing inclusion of items that generated variable performance across participants in the pilot phase with items sampled equally over the length of the movie.

2.2.3.3 Timeline judgements

This section followed the same procedure as described for piloting. Twenty-eight items were presented it total. Most of these items (24 of 28) were sourced from the piloted items (40) of this task, favouring those that elicited variable performance across participants over items that demonstrated floor or ceiling effects. Just as in the pilot phase, item selection was based on sampling items across the entire length of the movie, and sampling still frames separated by different interval lengths in the movie. In addition,

for the current phase, consideration was given to whether still frames sampled at shorter intervals apart (15 and 30 s) were sampled within or across scene changes. This latter consideration was important for Chapter 4, in which scene changes were conceptualized as event boundaries to test the impact of event boundaries on the temporal cohesion of memory. Scoring of the timeline judgements was also expanded for this phase. Three outcome measures were computed: position estimation error, interval estimation error, and correct ordering. Position estimation error, as in the piloting phase, reflected the average absolute deviation between the still frame's true position and the position estimated by the participant, expressed in seconds. Unlike the piloting phase, this measure was only averaged across still frames for items in which the two still frames were correctly ordered (i.e., the leftmost mouse click corresponded to the earlier still frame in the movie). Interval estimation error refers to the absolute difference between the true interval at which the still frames were sampled and the interval between mouse clicks (i.e., the judged interval), converted to seconds, and averaged across correctly ordered trials. Finally, correct ordering refers to the total number of items for which the two still frames were correctly ordered.

2.2.3.4 Comprehension questions

This section was unchanged from the pilot phase.

2.2.3.5 Neuropsychological tests

HC_{YA} participants completed the Logical Memory test and CALT as described above. They also completed another measure of verbal memory, the Rey Auditory Verbal Learning Test (RAVLT; Strauss, Sherman, & Spreen, 2006a), in which participants are asked to learn and later recall a list of 15 words. The delayed recall measure was used for analysis. The episodic subscale of the Survey of Autobiographical Memory (SAM; Palombo, Williams, Abdi, & Levine, 2013) was included as a self-report measure of naturalistic memory abilities. Finally, the Matrix Reasoning subtest of the Wechsler Adult Intelligence Scale-IV (WAIS-IV; Wechsler, 2008) was included as a measure of nonverbal/fluid intelligence.

2.3 Clinical validation

2.3.1 Rationale

The clinical-validation phase represented the culmination of revisions and refinements that were informed by the previous phases of data collection. The combined fMRI/memory protocol was completed in a sample of participants with refractory temporal lobe epilepsy (TLE) and demographically matched control participants, so that questions regarding the validity and sensitivity of the neuroimaging and memory measures could be addressed in the clinical population of interest. Data from this phase have been incorporated into Chapters 3 (movie memory), 4 (temporal memory for the movie), and 5 (neuroimaging and memory).

2.3.2 Participants

Participants with TLE were recruited from the Adult Epilepsy Service, University Hospital, London, who were undergoing presurgical evaluation for a temporal-lobe resection to control seizures. These individuals were contacted through a clinical research coordinator. In total, nineteen participants with TLE were enrolled (age $M \pm SD = 33.79 \pm 12.25$; years of education $M \pm SD = 12.79 \pm 2.25$; 9 female). A sample of neurologically healthy controls who were demographically matched to the TLE participants (referred to as HC_{TLE}) were recruited from the London community through flyers. Twenty-four control participants were enrolled (age $M \pm SD = 35.67 \pm 14.89$; years of education $M \pm SD = 13.74 \pm 1.76$; 14 female), all of whom self-identified as native English speakers with no history of hearing impairment, psychiatric illness, or neurological disorder. All participants provided informed consent and received nominal compensation to participate.

2.3.3 Procedure

Participants completed the same procedure as the healthy-validation sample, with some changes to the neuropsychological battery administered (see below). Notably, four

participants (3 clinical, 1 control) completed only the movie-driven fMRI procedure and not subsequent cognitive testing.

2.3.3.1 Free & probed recall

The procedure and scoring were identical to that used in the healthy-validation phase. The same two raters who conferred on the scoring guidelines applied to the HC_{YA} recall transcripts also independently scored a subset of the clinical-validation phase transcripts (10 TLE, 10 HC_{TLE}). The final inter-rater reliability coefficients for the healthy- and clinical-validation phases were used to guide the selection of outcome measures to be used for analysis.

2.3.3.2 Familiarity judgements

See healthy-validation phase.

2.3.3.3 Timeline judgements

See healthy-validation phase.

2.3.3.4 Comprehension questions

See healthy-validation phase.

2.3.3.5 Neuropsychological tests

For the TLE sample, the following scores were obtained from the clinical neuropsychological work-up: Logical Memory delayed recall from the Wechsler Memory Scale-Third Edition (story recall; Wechsler, 1997), California Verbal Learning Test-Second Edition (CVLT) delayed recall (word-list recall; Delis, Kaplan, Kramer, & Ober, 2000), Rey Complex Figure Test (RCFT) delayed recall (visual design memory; Strauss et al., 2006b) and Vocabulary and Matrix Reasoning from the Wechsler Abbreviated Scale of Intelligence-Second Edition (crystallized/verbal and fluid/nonverbal intelligence, respectively; Wechsler, 2011). Additional measures were obtained for both the TLE and HC_{TLE} samples: Rey Visual Design Learning Test (RVDLT) delayed recall (abstract

design-list recall; Spreen & Strauss, 1991), Names and Doors subtests from the Doors and People Test (verbal and visual recognition; Baddeley, Emslie, & Nimmo-Smith, 2006), and Conditional Associative Learning Test (CALT) trials to criterion (Petrides, 1985; St-Laurent, McCormick, et al., 2014). Finally, the HC_{TLE} sample was also administered Vocabulary & Matrix Reasoning subtests of the Wechsler Adult Intelligence Scale-Fourth Edition (crystallized/verbal and fluid/nonverbal intelligence, respectively; Wechsler, 2008).

Chapter 3

3 Movie memory: Novel stimuli for neuropsychological assessment in temporal lobe epilepsy

Neuropsychological assessment of episodic memory is a powerful tool to infer the functional integrity of the medial temporal lobes. The question of spared versus impaired medial temporal-lobe functioning is at the forefront of neuropsychological assessment in medically refractory temporal lobe epilepsy (TLE), when seizures are not well managed through pharmacological intervention. For these individuals, surgical resection of the epileptogenic tissue can be highly effective, with 70 to 90% of individuals expected to be free of disabling seizures following surgery (Engel, 2001). However, the benefit of seizure freedom must be weighed against the risk to sensory, motor, and/or cognitive abilities that may follow a cortical resection. Hence, individuals with refractory epilepsy must undergo extensive multidisciplinary investigations, including neuropsychological assessment, to ensure that they are good candidates for surgery and if they are deemed so, to establish a surgical plan.

Since the 1950s, when the importance of neuropsychological assessment in surgical planning for epilepsy was first recognized, the nature and scope of these assessments has evolved with advances in technology, medical practice, and clinical neuroscience (McAndrews & Cohn, 2012). However, despite ever-improving neuroimaging capabilities, neuropsychologists are still called upon to address questions regarding localization and lateralization of the seizure focus. The value of neuropsychology in detecting localizable impairments may be most apparent in non-lesional/cryptogenic cases of epilepsy, in which an epileptogenic lesion has not been identified on neuroimaging. However, even when localization hypotheses arise based on other investigations, the synthesis of different streams of evidence can enhance the interpretation of each and increase overall confidence in a proposed treatment plan (Baxendale & Thompson, 2010). Furthermore, presurgical neuropsychological performance is a strong predictor of postsurgical cognitive morbidity, especially when it

is integrated with other investigations (Baxendale, Thompson, Harkness, & Duncan, 2006). Memory decline postsurgically increases with better presurgical functioning of the hippocampus ipsilateral to the seizure focus (the functional adequacy hypothesis) and worse presurgical functioning of the contralateral hippocampus (the functional reserve hypothesis; Chelune, 1995). Thus, in combination with other standard investigations, neuropsychological assessment is a valuable tool in the localization of seizure foci and the prediction of postoperative decline.

In TLE, a foundational principle of neuropsychological assessment is the detection of material-specific effects on memory to lateralize medial temporal-lobe pathology. Assuming left-hemisphere dominance for language, left TLE is associated with impaired verbal memory and right TLE with nonverbal memory. This principle stems from some of the earliest neuropsychological investigations of epilepsy (particularly, postsurgical observations; Milner, 1972), and has continued to draw support from the literature (e.g., Gleissner et al., 1998; Glosser et al., 1995; Helmstaedter et al., 1991; Kim et al., 2003; Sass et al., 1995). Direct comparison of performance on verbal and nonverbal memory tests – ideally ones that are procedurally matched and co-normed on the same standardization sample (Jones-Gotman et al., 2010) – can shed light on the relative integrity of the left and right medial temporal regions. To increase specificity of material-dependent effects, memory stimuli that are simple and lend themselves to encoding in a single modality (verbal or nonverbal) are favoured over complex, multimodal stimuli that allow different strategies for successful encoding.

Although using a material-specificity framework can contribute important information for seizure lateralization and postsurgical prognosis, investigators commonly find that the link between visual memory and right TLE is harder to establish than the link between verbal memory and left TLE, even after unilateral temporal lobectomy (Baxendale & Thompson, 2010; Bell & Davies, 1998; Lee et al., 2002; McAndrews & Cohn, 2012; Saling, 2009). In his review of the topic, Saling (2009) noted that the material-specificity model is based on two flawed assumptions: (a) that verbal and nonverbal memory are

unitary constructs, and (b) that verbal memory exclusively relies upon left temporal regions, and non-verbal memory on right temporal regions. Adding to the uncertainty, individuals with epilepsy with early-life damage to the left temporal lobe are more likely to exhibit atypical (right or bilateral) hemispheric dominance for language compared to the general population (Branch et al., 1964; Duchowny et al., 1996; Hamberger & Cole, 2011; Möddel et al., 2009; Rasmussen & Milner, 1977; Rausch & Walsh, 1984; Springer et al., 1999) with potential concomitant reorganization of memory (Alessio et al., 2013; Gleissner, Helmstaedter, & Elger, 2002; Powell et al., 2007; Richardson, Strange, Duncan, & Dolan, 2003; Seidenberg et al., 1997). Therefore, inferring lateralized temporal-lobe damage from a unimodal memory impairment can be unreliable.

To complement the material-specificity assessment approach, another avenue to investigate functional integrity of the medial temporal lobe may be to explore natural, multimodal memory. In autobiographical memory, for instance, the hippocampus and other medial temporal structures play a clear role in retrieval (Svoboda, McKinnon, & Levine, 2006), and autobiographical-memory deficits are well-documented in TLE (Addis, Moscovitch, & McAndrews, 2007; Herfurth, Kasper, Schwarz, Stefan, & Pauli, 2010; McCormick, Moscovitch, Valiante, Cohn, & McAndrews, 2018; St-Laurent, Moscovitch, Levine, & McAndrews, 2009; Viskontas, McAndrews, & Moscovitch, 2000). However, as lab/clinic-based tools, autobiographical interviews have several limitations. First, they are difficult to verify. There are numerous reasons why an individual may produce details in an autobiographical interview that never occurred, but could nonetheless be counted toward his or her performance on the task. Second, autobiographical memories vary widely on factors that are expected to influence memory consolidation, such as the emotional salience of the memory or the complexity of events (Daselaar et al., 2008; St-Laurent, Moscovitch, et al., 2014).

Alternatively, perceptually rich and cognitively engaging movies would be expected to mimic the cognitive demands of everyday episodic memory while affording greater experimental control than autobiographical events (Furman et al., 2007). To our

knowledge, only one study has capitalized on the strengths of movie stimuli to investigate episodic-memory deficits in TLE. St-Laurent and colleagues (2014) investigated the effect of perceptual richness on the episodic recall of participants with TLE. Participants were exposed to a set of perceptually enriched (audiovisual film clips) and impoverished (written narratives) events matched on story content, and were later asked to recall story content and perceptual details. Compared to controls, participants with TLE recalled fewer details overall but disproportionately fewer perceptual details. The researchers also conducted an autobiographical-memory task for comparison and found the same relative discrepancy in recall of perceptual details. Converging evidence across their autobiographical-memory and lab-based event-memory paradigms suggest that both tasks were sensitive to a general lack of vividness in the episodic recall of their TLE sample.

In the present study, we sought to design a memory test that was sensitive to medial temporal-lobe dysfunction in TLE that might complement the routine neuropsychological assessment of memory. To that end, we selected a brief, suspenseful audiovisual film clip for our stimulus that has previously been used to explore naturalistic brain activation in cognitive neuroscience research (Ben-Yakov & Henson, 2018; K. L. Campbell et al., 2015; Hasson et al., 2010; Naci, Cusack, Anello, & Owen, 2014; Taylor et al., 2017). The test itself was composed of different question types, described below, that were designed to maximize sensitivity to medial temporal-lobe dysfunction. Data were collected in three phases. A pilot phase with healthy young adults was conducted to refine test items and evaluate construct validity in a healthy sample. Next, the healthy-validation phase with a separate group of healthy young adults was conducted to refine scoring procedures and measures of interest. Finally, the clinical-validation phase included a sample of TLE participants and demographically matched controls to further investigate properties of the test in our clinical sample. Our aim was to describe some of the psychometric properties of this novel test as well as its sensitivity to TLE.

3.1 Methods

3.1.1 Movie stimulus

The movie stimulus was sampled from a 1961 television episode entitled, “Alfred Hitchcock Presents: Bang! You’re Dead.” The original 20-minute episode was edited down to eight minutes while maintaining the narrative structure of the original. This black-and-white audiovisual clip depicts a boy playing with a real gun that he believes to be a toy. This stimulus has been shown to reliably activate large parts of the cerebral cortex (Hasson et al., 2010) and its suspenseful plot promotes engagement with the task (Naci et al., 2014). It was also novel to most viewers – all participants in the current study reported not having seen it before, except one control participant, who was unsure – and it is, therefore, more likely to robustly engage medial temporal-lobe structures (Kumaran & Maguire, 2009; Tulving, Markowitsch, Craik, Habib, & Houle, 1996).

3.1.2 Test construction

Different question types were employed to investigate different aspects of memory for the movie stimulus. Since the movie was multimodal and can be encoded using visual and/or verbal strategies, different question types were expected to facilitate visual, verbal, or combined visual and verbal retrieval strategies. The free and probed recall section was completed first as an oral interview, and then the familiarity judgements and general comprehension questions were administered on laptop, using the Psychophysics toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) in MATLAB 2014b (Mathworks, Natick, MA, USA). Sample items for the computer-administered sections are shown in Figure 2. The structured interview script and complete list of comprehension questions are provided in Appendices A and D. A fourth question type for which participants were asked to make temporal judgements about when events occurred in the context of the movie was investigated separately (see Chapter 4) and will not be revisited here.

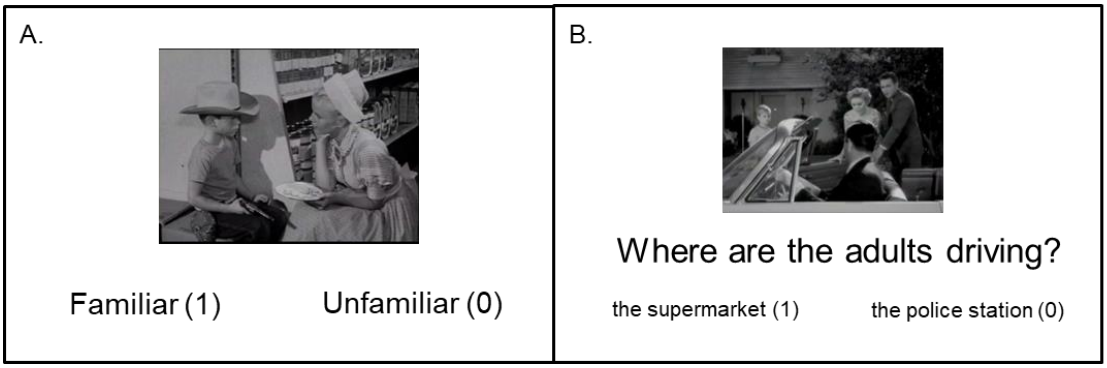


Figure 2. Sample computer-administered items for the familiarity judgements (*A*) and general comprehension questions (*B*) from the movie-memory test.

Free and probed recall was a structured interview modeled after the Autobiographical Interview (AI; Levine et al., 2002). The AI has been shown to be sensitive to medial temporal dysfunction in TLE (Herfurth et al., 2010; Park, St-Laurent, McAndrews, & Moscovitch, 2011; Sheldon, McAndrews, & Moscovitch, 2011; St-Laurent et al., 2009, 2011). In one study, the protocol was adapted to probe memory for brief movies and demonstrated sensitivity to deficits in episodic recall (St-Laurent, Moscovitch, et al., 2014). The adapted interview followed a staged cueing procedure composed of a free-recall condition, a general probe, and 15 specific probes. In free recall, participants were asked to recall the movie in as much detail as possible. The general probe prompted participants to provide any additional information that they had not recounted in free recall. Specific probes gave participants a final opportunity to recall details of the movie pertaining to when the events took place (e.g., the time of day), where the events took place (e.g., settings/buildings), the main characters, perceptual details (e.g., non-speech sounds, room decor), as well as thoughts and emotions experienced while observing the movie.

In *familiarity judgements*, participants were shown a still frame and asked to judge whether it was familiar or unfamiliar. A 1:1 target:lure ratio was used. Target items were sampled from the eight-minute movie clip shown to participants. An equal number of items was sampled from each of the two-minute quarters of the movie to ensure adequate

sampling from the entire length of the movie. Lure items were sampled from the unused 12 minutes of the original television episode. In lure items, the settings and characters were likely to have been depicted in the eight-minute clip (i.e., within the original 20 minutes, they were offset by seconds or minutes from the content used for the eight-minute clip), so participants had to be sensitive to whether the whole scene was one they had viewed in the movie clip. The order of items was pseudorandomized such that there were no more than three items of either type in a row. Familiarity is thought to be subserved by medial temporal areas such as perirhinal cortex, and may not specifically rely on hippocampal functioning (Bowles et al., 2007; Ranganath et al., 2004). However, the high degree of similarity between the targets and lures may tax the pattern-separation abilities of the hippocampus (Yassa & Stark, 2011).

The *general comprehension questions* were constructed to assess whether the participant was following the plot of the movie. The questions took a two-option forced choice format and, when appropriate, were accompanied by a still frame from the movie to provide contextual support. Some of the questions could only be answered by recalling specific scenes that were on screen for a few seconds. These were sampled from the full length of the movie. For example, the question “What are the boys pretending to shoot at?”, accompanied by a picture of the first scene in which two boys are standing behind a tree, would require the participant to recall the first 20 seconds of the movie in which the two boys are pretending to exchange fire with a third boy standing at a distance. Other questions could be answered by recalling one occurrence of a repeated event (e.g., “What type of mechanical animal does the boy ride?” when the boy is shown to be riding in multiple scenes) or integrating information across scenes (e.g., “In total, how many bullets did the boy put in the gun?”).

3.1.3 Pilot testing & test revision

A preliminary version of the memory test was piloted in 30 neurologically healthy young adults (21 female), recruited from the local undergraduate population (age $M = 18.63$, $SD = 0.96$). The purpose of the pilot was to provide preliminary evidence of construct

validity (through correlations with standardized memory tests) and to refine the familiarity-judgement items.

Participants completed the free- and probed-recall interview, familiarity-judgement task (24 items), and general comprehension questions (20 questions). To score free and probed recall, a checklist of events and other detail types was devised from the plot of the movie. For each participant, the number of correctly reported details in the checklist were tallied. For familiarity judgements, a d′ score was computed. To evaluate a larger sample of familiarity-judgement items, 15 participants were randomly assigned to each of two versions of the familiarity-judgement task with no overlapping items. For general comprehension questions, the total number of correct responses was tallied. In addition to the movie-memory test, participants were asked to complete a brief battery of standardized neuropsychological measures: the Logical Memory subtest of the Wechsler Memory Scale-Fourth Edition (story recall; Wechsler, 2009), the Rey Complex Figure Test (RCFT; visual recall; Strauss et al., 2006b), the Conditional Associative Learning Test (CALT; spatial associative learning; Petrides, 1985), and the Digit Span subtest of the Wechsler Adult Intelligence Scale-Fourth Edition (auditory attention; Wechsler, 2008). The movie-memory measures were expected to correlate most strongly with Logical Memory (as a measure of recall for contextualized information), followed by other memory tests (RCFT, CALT), and least strongly or not at all with Digit Span.

Correlations with standardized cognitive measures (under false discovery rate, FDR, correction; Benjamini & Hochberg, 1995) revealed significant correlations between Logical Memory delayed recall and total details recalled in free and probed recall (r = 0.55, p = .002) as well as accuracy on the general comprehension questions (r_s = 0.48, p = .007), but the correlation with familiarity-judgement d′ did not survive FDR correction. No other significant correlations were detected. These findings suggest that the movie memory questions show convergent validity with a standardized measure of contextualized verbal (i.e., story) memory.

To refine the familiarity-judgement items, the proportion of participants that successfully judged each item was determined. Items for which eight to 13 of 15 participants made a correct judgement were favoured over items that were consistently difficult (i.e., fewer than eight of 15 participants correctly judged the item, signifying below chance responding) or consistently easy (i.e., 14 to 15 of 15 participants correctly responded). A set of 16 items was selected, balancing inclusion of items that generated variable performance across participants with items sampled equally over the length of the movie.

3.1.4 Healthy- & clinical-validation phases

3.1.4.1 Participants

For a summary of demographic and other sample characteristics, see Table 3. Twenty-four participants (referred to as the HC_{YA} sample) were recruited for the healthy-validation phase of data collection. All participants reported no history of neurological or psychiatric illness and were native English speakers (learned English before age 5), with self-reported normal hearing and normal or corrected-to-normal vision.

The clinical-validation phase consisted of a TLE sample (n = 16) and a healthy control sample (HC_{TLE}; n = 22). Participants with TLE were recruited from the Adult Epilepsy Service at London Health Sciences Centre (University Hospital) in London, Ontario, Canada. All were identified as potential surgical candidates for a temporal-lobe resection to control seizures, based on electroencephalography (EEG) evidence of TLE, corroborated in some cases by a structural lesion in the temporal lobe on magnetic resonance imaging (MRI). Clinical MRI was read as showing probable mesial temporal sclerosis (MTS) in eight participants, possible MTS in two participants, a non-MTS structural abnormality in two participants, and no structural abnormality in four participants. None had undergone previous brain surgery. The inclusion criteria for the HC_{TLE} sample were the same as those used for the HC_{YA} sample. However, the HC_{TLE} sample was matched to the TLE sample on age (two-tailed independent-groups t-test; p = .988), education (p = .068), and sex distribution (p = .782).

Table 3. Participant characteristics and cognitive abilities.

	HC_{YA}	HC_{TLE}	TLE
n	24	22	16
Sex (F:M)	14:10	12:10	8:8
Age (*M* ± *SD*)	23.17 ± 3.24	35.18 ± 15.41	33.88 ± 11.85
Years of education (*M* ± *SD*)	15.63 ± 2.32	13.43 ± 1.47	12.56 ± 2.28
Handedness (R:L)	24:0	17:5	16:0
Seizure lateralization	.	.	6 L; 9 R; 1 BL
Evidence of MTS on MRI			8 probable; 2 possible; 6 none
Years since onset (*M* ± *SD*)	.	.	19.07 ± 17.66
SAM Episodic	98.68 ± 11.01	.	.
Vocabulary (scaled)	.	12.45 ± 2.74	8.60 ± 2.72
Matrix Reasoning (scaled)	12.58 ± 2.15	11.00 ± 2.66	9.33 ± 2.38
Logical Memory (scaled)	11.17 ± 1.97	.	7.60 ± 3.96
CVLT	.	.	8.27 ± 4.74
RAVLT	12.13 ± 2.17	.	.
RVDLT	.	11.82 ± 2.75	10.06 ± 3.28
Names	.	19.95 ± 2.73	18.69 ± 2.60
RCFT	.	.	13.93 ± 4.95
Doors	.	20.91 ± 2.47	17.27 ± 4.03
CALT	29.79 ± 14.11	35.62 ± 18.62	45.75 ± 18.78

Note. HC_{YA} = healthy control participants – young adults; HC_{TLE} = healthy control participants – matched to epilepsy sample; TLE = temporal lobe epilepsy sample; F = female; M = male; R = right; L = left; BL = bilateral; MTS = mesial temporal sclerosis; MRI = magnetic resonance imaging; SAM = Survey of Autobiographical Memory; CVLT = California Verbal Learning Test; RAVLT = Rey Auditory Verbal Learning Test; RVDLT = Rey Visual Design Learning Test; RCFT = Rey Complex Figure Test; CALT = Conditional Associative Learning Test.

3.1.4.2 Movie-memory test

Study participants viewed the movie while undergoing a functional MRI scan (the results of which are not reported here). Approximately 10 minutes passed between the end of the movie and initiation of the movie-memory task. Participants were not forewarned that their memory for the movie would be tested; they were simply instructed to watch the clip as they would normally watch a television episode or movie and to follow the plot as best they can.

All three study samples received the same version of the movie-memory test, including the free- and probed-recall interview, familiarity judgements (16 items), and comprehension questions (20 items). Scoring for the familiarity judgements and comprehension questions was consistent with the pilot phase; response sensitivity for familiarity judgements was captured with d′ and accuracy on the general comprehension questions was measured as the sum of the correct responses. For these phases, a more sophisticating scoring system for the free- and probed-recall section was devised. Just as the interview had been adapted from the AI, the new scoring system was also modeled after the AI scoring guidelines (see Appendix B).

Audio recordings of the recall interview were transcribed. Transcripts in the clinical-validation phase were deidentified to conceal the study sample (TLE or HC_{TLE}) to which the participant belonged. Transcripts were first segmented into discrete details and categorized as internal or external. *Internal details* referred to any information that pertained to the content of the movie (e.g., subevents, perceptual details, etc.) or the participant's experience of viewing the movie (e.g., thoughts and emotions experienced during movie viewing). *External details* did not specifically pertain to the movie or represented extraneous or repeated information. Internal details were also classified as correct or incorrect. A detail was labeled as incorrect when it opposed information available in the movie. Details that could not be verified (e.g., a participant's report of their own emotions) or those that were ambiguous were coded as correct. The primary

outcome measure for the free and probed recall section was the total number of correct internal details recalled.

Internal and external details were further classified into eight exclusive detail types: (1) *event details* (referring to the unfolding of events, introduction of characters, character dialogue); (2) *place details* (referring to buildings, rooms, locations within a room); (3) *time details* (referring to the era, season, time of day); (4) *perceptual details* (referring to visual and auditory details); (5) *emotion/thought details* (referring to what the participant was thinking and/or feeling while viewing the movie); (6) *semantic details* (referring to general knowledge); (7) *repetitions*; and (8) *other details* (a category of exclusion). Types 1 to 5 could be internal or external whereas types 6 to 8 were necessarily external. See Figure 3 for a sample scored transcript.

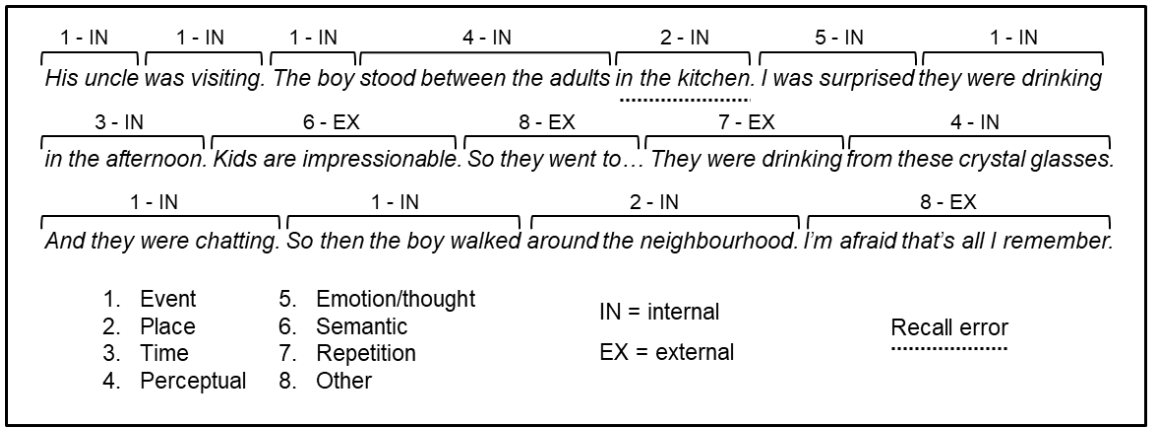


Figure 3. A sample transcript of a recall response, demonstrating segmentation and characterization of details and errors.

The theoretical distinction between these detail types may be more tenuous for recall of a movie compared to recall of an autobiographical event. Informed by the scoring methods of St-Laurent and colleagues (2014) who adapted the AI for a brief movie stimulus, we also subclassified internal details as *story details* – combining event details and emotion/thought details, seen as mental events – and *perceptual details* – including

perceptual details (as defined above) and time and place details, as these would be inferred from perceptual information in the movie.

Two independent raters scored each of the 24 recall transcripts of the healthy-validation phase. Inter-rater reliability was evaluated for recall measures, and the results were used to refine the variables included in analysis. Intraclass correlation coefficients were calculated based on a single-rating, absolute-agreement, two-way mixed effects model, a conservative model that provides the most appropriate estimate of reliability, assuming one clinician/researcher as sole rater (McGraw & Wong, 1996). Data from the healthy-validation phase was assessed first to refine the scoring guidelines. After a discussion of scoring discrepancies between raters, the scoring guidelines were revised and the transcripts were rescored. The revised guidelines were also used by the same two raters to independently rate a set of 20 (10 TLE) transcripts of the clinical-validation phase. These three sets of inter-rater reliability coefficients are displayed in Table 4. Several detail types were associated with poor inter-rater reliability, likely a consequence of low frequencies of the detail type. Summary statistics that combined different detail types (internal, external, story, perceptual) typically showed greater reliability. These summary scores, based on the ratings of a single scorer who rated the remaining 18 transcripts, were the scores used for analysis.

Table 4. Inter-rater reliability for recall transcript measures.

Detail type	Healthy Validation Before revising scoring guidelines	Healthy Validation After revising scoring guidelines	Clinical Validation
Total Internal	**0.67**	**0.70**	**0.77**
Event	0.68	0.68	0.78
Emotion/thought	0.70	0.76	0.92
Total Story	**0.67**	**0.66**	**0.80**
Place	0.74	0.78	0.86
Time	0.06	0.29	0.66
Perceptual	0.69	0.78	0.67
Total Perceptual	**0.69**	**0.79**	**0.72**
Total External	**0.83**	**0.83**	**0.83**
Semantic	0.39	0.47	0.57
Repetitions	0.90	0.85	0.75
Other	0.79	0.74	0.88

Note. Variables selected for further analysis are shown in **bold.**

3.1.4.3 Standardized cognitive battery

A number of standardized cognitive measures were also administered (see Table 3). For the HC_{YA} sample, we obtained Logical Memory delayed recall from the Wechsler Memory Scale-Fourth Edition (story recall; Wechsler, 2009), Rey Auditory Verbal Learning Test (RAVLT) delayed recall (word-list recall; Strauss et al., 2006a), Conditional Associative Learning Test (CALT) trials to criterion (spatial associative learning; Petrides, 1985; St-Laurent, McCormick, et al., 2014), Survey of Autobiographical Memory (SAM) episodic memory subscale score (self-report episodic memory; Palombo et al., 2013), Matrix Reasoning subtest of the Wechsler Adult Intelligence Scale-Fourth Edition (fluid/nonverbal intelligence; Wechsler, 2008).

For the TLE sample, the following scores were obtained from the clinical neuropsychological work-up: Logical Memory delayed recall from the Wechsler Memory Scale-Third Edition (story recall; Wechsler, 1997), California Verbal Learning Test-

Second Edition (CVLT) delayed recall (word-list recall; Delis et al., 2000), Rey Complex Figure Test (RCFT) delayed recall (visual design memory; Strauss et al., 2006b) and Vocabulary and Matrix Reasoning from the Wechsler Abbreviated Scale of Intelligence-Second Edition (crystallized/verbal and fluid/nonverbal intelligence, respectively; Wechsler, 2011). Additional measures were obtained for both the TLE and HC_{TLE} samples: Rey Visual Design Learning Test (RVDLT) delayed recall (abstract design-list recall; Spreen & Strauss, 1991), Names and Doors subtests from the Doors and People Test (verbal and visual recognition; Baddeley et al., 2006), and Conditional Associative Learning Test (CALT) trials to criterion (Petrides, 1985; St-Laurent, McCormick, et al., 2014). Finally, the HC_{TLE} sample was also administered Vocabulary & Matrix Reasoning subtests of the Wechsler Adult Intelligence Scale-Fourth Edition (crystallized/verbal and fluid/nonverbal intelligence, respectively; Wechsler, 2008). Just as in the pilot phase, the strongest correlations between movie-memory and standardized measures were expected with the Logical Memory test, as the most contextually rich memory test available, with weaker correlations expected with other memory tests, and little or no correlation with measures of intellectual function.

Raw scores were used in analysis, with the exception of Logical Memory, Vocabulary, and Matrix Reasoning, for which age-scaled scores were used to facilitate comparison across the two versions of these tests administered to different study samples (i.e., Wechsler Memory Scale-Third Edition in the TLE sample but Fourth Edition in the HC_{YA} sample, Wechsler Adult Intelligence Scale-Fourth Edition in HC_{YA} and HC_{TLE} samples but Wechsler Abbreviated Scale of Intelligence-Second Edition in the TLE sample).

3.2 Results

3.2.1 Construct validity

To investigate construct validity, we performed correlations across movie-memory and standardized cognitive measures. The three study samples (HC_{YA}, HC_{TLE}, TLE) were pooled on measures that were administered to more than one sample to improve

statistical power. All bivariate correlations across movie-memory and standardized cognitive measures are shown in Table 5. Within this large correlation matrix, several significant correlations survive FDR correction. Recall of internal details significantly correlated with Logical Memory (delayed story recall; r_s = 0.48, p = .002), Doors (visual recognition; r_s = 0.46, p = .004), and Vocabulary (verbal intelligence; r_s = 0.46, p = .005). Familiarity d′ significantly correlated with Doors (r_s = 0.47, p = .003). Comprehension accuracy significantly correlated with Doors (r_s = 0.58, p < .001) and Vocabulary (r_s = 0.43, p = .008). For context, correlations across standardized cognitive measures are provided in Table 6.

Table 5. Correlations across movie-memory and standardized cognitive measures.

		Recall Internal Details		Familiarity d′		Comprehension Accuracy	
	n	*r*	*p*	*r*	*p*	*r*	*p*
Logical Memory	39	0.48**	.002	0.26	.108	0.25	.126
RCFT	15	0.37	.178	0.38	.160	0.13	.636
CVLT	15	0.56*	.031	0.17	.545	-0.08	.772
RAVLT	24	0.24	.265	0.46*	.025	0.12	.589
RVDLT	38	0.21	.214	0.37*	.021	0.37*	.022
Names	38	0.02	.911	0.27	.096	0.23	.164
Doors	37	0.46**	.004	0.47**	.003	0.58**	.000
CALT	61	-0.07	.595	-0.20	.124	0.04	.775
Vocabulary	37	0.46**	.005	0.27	.107	0.43**	.008
Matrix Reasoning	61	0.01	.935	0.23	.074	0.10	.444
SAM Episodic	24	-0.21	.318	-0.19	.367	-0.14	.502

Note. Correlations are collapsed across study samples for tests administered to more than one sample. RCFT = Rey Complex Figure Test; CVLT = California Verbal Learning Test; RAVLT = Rey Auditory Verbal Learning Test; RVDLT = Rey Visual Design Learning Test; CALT = Conditional Associative Learning Test; SAM = Survey of Autobiographical Memory.
* significant at the .05 level (but do not survive false discovery rate correction)
****** significant after controlling the false discovery rate

To gain a deeper understanding of these significant correlations, they were statistically compared across samples. We computed 95% confidence intervals (CI) for the difference between the correlations, where an interval that includes 0 suggests a nonsignificant

difference (Zou, 2007). Among the significant correlations in the pooled sample, the correlation between comprehension question accuracy and Doors was found to significantly differ across its constituent samples (TLE and HC_{TLE}), 95% CI: [0.06,1.21]. A strong correlation was detected in the HC_{TLE} sample ($r_s = 0.67$, $p = .001$) but not in the TLE sample ($r < 0.01$, $p = .995$), suggesting that the construct measured by the comprehension questions may differ in the HC_{TLE} and TLE groups.

Table 6. Correlations across standardized cognitive measures.

	Logical Memory			RCFT			CVLT			RAVLT			RVDLT		
	n	*r*	*p*	*n*	*r*	*p*	*n*	*r*	*p*	*n*	*r*	*p*	*n*	*r*	*p*
RCFT	15	0.72**	.003												
CVLT	15	0.76**	.001	15	0.59**	.021									
RAVLT	24	0.07	.759	0			0								
RVDLT	15	0.47	.074	15	0.68**	.005	15	0.48	.069	0					
Names	15	0.26	.345	15	0.26	.348	15	0.45	.094	0			38	0.54**	.000
Doors	14	0.62**	.018	14	0.72**	.004	14	0.41	.151	0			37	0.57**	.000
CALT	39	-0.22	.186	15	-0.29	.293	15	-0.35	.200	24	-0.15	.484	37	-0.47**	.003
Vocabulary	15	0.24	.380	15	0.48	.070	15	0.15	.591	0			37	0.29	.078
Mat. Reasoning	39	0.41**	.009	15	0.58*	.024	15	0.27	.338	24	0.31	.134	37	0.50**	.001
SAM Episodic	24	0.05	.833	0			0			24	-0.06	.777	0		

	Names			Doors			CALT			Vocabulary			Matrix Reasoning		
	n	*r*	*p*	*n*	*r*	*p*	*n*	*r*	*p*	*n*	*r*	*p*	*n*	*r*	*p*
RCFT															
CVLT															
RAVLT															
RVDLT															
Names															
Doors	37	0.43**	.008												
CALT	37	-0.51**	.001	36	-0.47**	.004									
Vocabulary	37	0.30	.071	36	0.54**	.001	36	-0.36*	.032						
Mat. Reasoning	37	0.48**	.003	36	0.56**	.000	60	-0.48**	.000	37	0.46**	.004			
SAM Episodic	0			0			24	-0.13	.535	0			24	0.04	.853

Note. Correlations are collapsed across study samples for tests administered to more than one sample. RCFT = Rey Complex Figure Test; CVLT = California Verbal Learning Test; RAVLT = Rey Auditory Verbal Learning Test; RVDLT = Rey Visual Design Learning Test; CALT = Conditional Associative Learning Test; SAM = Survey of Autobiographical Memory.
* significant at the .05 level (but do not survive false discovery rate correction)
****** significant after controlling the false discovery rate

3.2.2 Group differences

Performance on the movie memory test was compared across TLE and HC_{TLE} samples. First, a mixed-model ANOVA was conducted to test the effect of detail type (internal vs. external) and group on the number of details recalled during free and probed recall (see Figure 4A). A significant main effect of group ($F(1,36) = 4.27$, $p = .046$, $\eta^2 = .11$) indicated that control participants produced significantly more details overall. A significant main effect of detail type ($F(1,36) = 232.29$, $p < .001$, $\eta^2 = .83$) showed that internal details were more commonly reported than external details. There was also a significant interaction effect of group and detail type ($F(1,36) = 12.81$, $p = .001$, $\eta^2 = .05$). Follow-up simple effect analyses revealed that, compared to TLE participants, control participants produced significantly more internal details ($F(1,36) = 8.34$, $p = .007$, $\eta^2 = .19$) but a statistically comparable number of external details ($p = .553$).

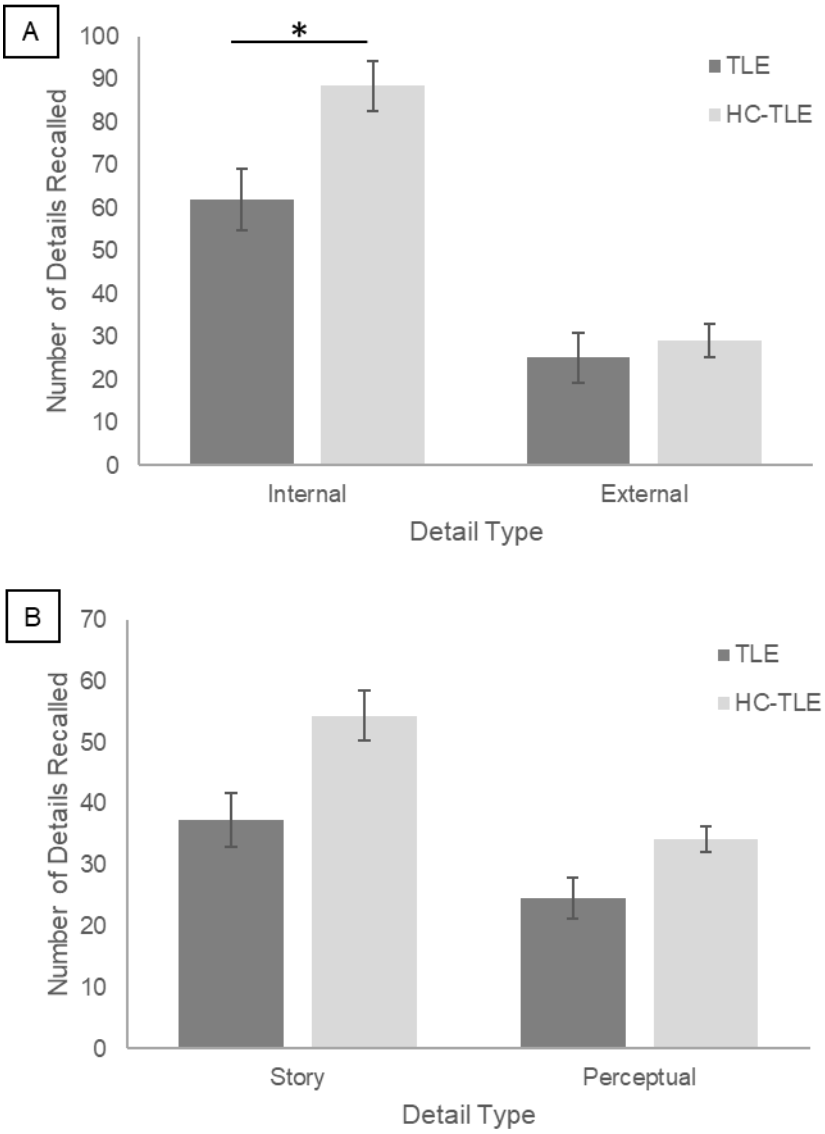


Figure 4. Group comparisons by detail type. *A* depicts the number of internal and external details recalled by group, and *B* depicts the of story and perceptual details recalled by group. Error bars represent standard error of the mean. TLE = temporal lobe epilepsy sample; HC-TLE = healthy control participants – matched to epilepsy sample.

* significant at the .05 level

A second mixed-model ANOVA comparing group performance on recall of correct internal detail subtypes – story and perceptual details – demonstrated a significant main effect of group (HC_{TLE} < TLE; $F(1,36) = 8.34$, $p = .007$, $\eta^2 = .19$) and a significant main effect of detail subtype (story > perceptual; $F(1,36) = 64.35$, $p < .001$, $\eta^2 = .62$). The effect of detail subtype did not differ between the groups ($F(1,36) = 3.21$, $p = .082$, $\eta^2 = .03$). These results are displayed in Figure 4B.

Outcome measures of the familiarity and comprehension subsections also significantly distinguish the TLE from the HC_{TLE} group (familiarity d′: $U = 90.0$, $p = .010$, $\eta^2 = .18$; comprehension accuracy: $U = 102.0$, $p = .029$, $\eta^2 = .13$), with control participants outperforming participants with TLE.

Examining movie-memory outcomes by laterality of epilepsy – left ($n = 6$) or right ($n = 9$) – revealed no difference between subgroups in recall of internal details ($p = .324$), familiarity d′ ($p = .999$), or general comprehension accuracy ($p = .500$). Similarly, comparing TLE participants with evidence of possible or probable MTS on structural MRI ($n = 10$) versus no evidence of MTS ($n = 6$) revealed no significant difference on recall of internal details ($p = .635$), familiarity d′ ($p = .220$), or general comprehension accuracy ($p = .447$).

3.2.3 Predicting group membership

Binary logistic regression was used to investigate the ability of movie-memory measures to predict group membership (TLE vs. HC_{TLE}). The main outcome measure of each of the three test sections were included in the model: recall of internal details, familiarity d′, and comprehension accuracy. Without a theory-driven framework to inform model building, a backward stepwise method was employed using the likelihood ratio statistic to guide exclusion of predictors. The original model was statistically significant ($\chi^2(3) = 15.86$, $p = .001$), explained 45% of the variance in group, and correctly classified 82% of cases. Comprehension accuracy was a nonsignificant predictor of group ($p = .427$) and was eliminated from the model. The final model was significant ($\chi^2(2) = 15.22$, $p < .001$),

explained 44% of the variance in group, and correctly classified 80% of cases. Recall of internal details and familiarity d′ were retained in the final model as significant predictors (recall: $B = 0.04$, Wald $\chi^2 = 4.98$, $p = .026$; familiarity: $B = 0.68$, Wald $\chi^2 = 4.47$, $p = .035$). For both predictors, higher performance was associated with decreased likelihood of having epilepsy (recall: *Exp(B)* = 1.04, 95% CI: [1.00, 1.07]; familiarity: *Exp(B)* = 1.97, 95% CI: [1.05, 3.71]). The sensitivity for this classification (i.e., proportion of TLE participants classified as TLE) was 0.69 and the specificity (i.e., proportion of HC_{TLE} participants classified as HC_{TLE}) was 0.87.

3.3 Discussion

3.3.1 Staged approach

In this study, we demonstrated preliminary evidence of inter-rater reliability, construct validity, and sensitivity to TLE for a movie-memory test. Data were collected in three phases: a pilot phase, in which an early version of the memory test was administered to a sample of healthy young adults ($n = 30$); a healthy-validation phase, in which the revised memory test was administered to a separate group of healthy young adults ($n = 24$); and a clinical-validation phase, consisting of a group of TLE participants ($n = 16$) and matched control participants ($n = 22$), in which analyses were refined based on the previous phases of data collection.

The pilot phase was conducted to investigate construct validity of a test prototype in a sample of healthy young adults and to refine test items to avoid ceiling and floor effects. Two of three movie-memory measures (free and probed recall and general comprehension questions) were found to strongly correlate with delayed recall on a standardized measure of story memory (Logical Memory; demonstrating convergent validity). Thus, we can be reasonably confident that, at least in this healthy sample, performance on the movie-memory test is tapping into a similar construct(s) related to contextualized episodic memory as this standardized measure of story recall.

In the healthy-validation phase, scoring of the free- and probed-recall interview was investigated for inter-rater reliability across two independent raters. The reliability results were used to refine the scoring guidelines and to identify measures with adequate reliability for further analysis. Revising the scoring guidelines led to some modest improvements in inter-rater reliability when the healthy-validation sample was rescored, and when the clinical-validation samples were scored for the first time. The measures selected for analysis were summary measures – internal, external, story, and perceptual details – rather than the tallies of each individual detail type, since the summary measures were shown to have "good" to "excellent" reliability based on conventional guidelines (Cicchetti, 1994). Furthermore, the grouping of detail types for these summary measures was theoretically motivated and mirrors previous relevant literature (St-Laurent, Moscovitch, et al., 2014).

3.3.2 Construct validity

Correlations between outcome measures of the movie-memory test and standardized measures of memory were conducted to elucidate the degree of overlap in their underlying constructs. Based on the novelty of these measures and their intended application, we did not seek to rigorously identify the constructs that are shared or not shared with standardized measures (D. Campbell & Fiske, 1959), but rather, at a more basic level, to explore the relationships between our novel measures and a variety of previously validated memory measures. The three samples with comparable datasets (i.e., those that were administered the revised memory test, and whose recall transcripts were scored with the revised guidelines) – the HC_{YA}, HC_{TLE}, and TLE samples – were pooled to improve statistical power. With respect to the free- and probed-recall interview, the total number of correct internal details was found to correlate with a measure of delayed story recall (Logical Memory), a measure of visual recognition memory (Doors), and a measure of verbal/crystallized intelligence (Vocabulary). As above, the link between details recalled for the movie and details recalled for a short story is not surprising. A correlation with Vocabulary, a test in which participants are asked to provide definitions for words, may suggest that word knowledge while viewing the movie or expressive

language abilities while recalling the movie can influence performance on the memory test, irrespective of memory abilities.

Sensitivity of responding on the familiarity-judgement task (measured by d′) was correlated with performance on the Doors test. In the Doors test, the participant sees a series of doors and is then shown a page with four doors on it and is asked to pick the door he or she saw earlier. The task involves two conditions, one in which the lures on the four-choice page are dissimilar to the target door, and one in which the lures are highly similar. This subtest is sensitive to visual-memory deficits following right temporal lobectomy (Morris, Abrahams, Baddeley, & Polkey, 1995). The familiarity-judgement task is structured in much the same way as the Doors test, in that a forced-choice paradigm is used to test recognition memory for a visual stimulus. As well, lure items in this task were sampled from the 12 minutes of the television episode that were not shown to the participant. Some were dissimilar to the target items (e.g., depicted new characters, scenes, or settings that were not shown in the eight-minute movie) and others were highly similar (e.g., sampled from the same scene as other target items). However, unlike the Doors test, encoding of the information in the movie was more amenable to a multimodal (verbal and non-verbal) strategy: participants had to choose between two options instead of four, and lures that were highly similar to target items were not administered alongside the target items to which they were similar.

Consistent with other movie-memory measures, accuracy on the comprehension questions was correlated with visual recognition memory (Doors) and verbal/crystallized intelligence (Vocabulary). Notably, when the strength of correlations detected in the pooled sample were compared across individual samples, the correlation between comprehension accuracy and the Doors tests was found to be significantly different in the TLE group compared to their healthy counterparts (HC_{TLE}). Specifically, the correlation was only significant in the control sample. This observation highlights an important consideration in the validation of novel assessment tools: that the tools may not measure the same constructs or respond to the same influences in healthy and clinical samples.

Here, we are interested in using the available data to begin to understand the underlying cognitive processes that contribute to these novel measures. Ultimately, in a larger sample of TLE participants, we would seek to explore how specific neural and cognitive changes documented in TLE influence test performance. Failure to detect a significant difference across groups in other correlations between movie and standardized memory measures does not rule out the possibility that different cognitive factors are at play in the performance of different subgroups, but it does support the pooling of groups to investigate these relationships.

3.3.3 Impairment in TLE

In group comparisons on movie-memory measures, the TLE group performed more poorly across the three test sections compared to healthy matched controls. Based on the magnitude of effect sizes (η^2), group differences accounted for 13% of the variance in comprehension accuracy, 18% of the variance in familiarity-judgement sensitivity, and 19% of the variance in recall of internal details. Internal details reported during free and probed recall were also subdivided into story details (pertaining to the unfolding of events in the story or "mental events" in the viewer's mind) and perceptual details (pertaining to auditory and visual details or details inferred from auditory/visual details). This analysis was informed by the study of St-Laurent and colleagues (2014) who found that participants with TLE produced fewer story and perceptual details than control participants, particularly fewer perceptual details. This relative disadvantage in vividness of recall was detected for an autobiographical-memory paradigm, but also when memory for brief audiovisual clips was assessed. In the present study, no interaction between group and internal detail subtype was detected: the TLE group recalled fewer details than the matched control group, with a similar relative disadvantage recalling story details as recalling perceptual details.

It is unclear why the relative vividness disadvantage was not detected here. There are numerous differences between the present study and that of St-Laurent and colleagues. For one, qualities of the movie stimuli themselves were starkly different, including the

length of the clips (23 seconds versus 8 minutes) and the verbal content of the clips (minimal or no dialogue versus near-continuous dialogue). It is possible that a longer stimulus and more dialogue provide more story details (disproportionate to the increase in perceptual details) that the healthy control participants could continue to learn and later recall, but that far exceeded the memory capacity of TLE participants. In this scenario, a greater story detail disadvantage in the present study might then be large enough to match a perceptual detail disadvantage. In addition, the composition of the TLE groups differed. St-Laurent and colleagues included TLE participants who were pre- or post-temporal lobectomy. Although they did not find a significant effect of surgery status on performance, it is still possible that inclusion of postsurgical participants affected the pattern of group differences on different recall measures. These ideas are simply speculation, and further research into the factors that promote or inhibit recall of story versus perceptual details may be valuable in optimizing assessment tools in TLE.

Movie-memory performance was also investigated in TLE subgroups based on laterality of seizure focus and evidence of MTS on MRI. No significant differences in performance were detected between left and right TLE subgroups or between those with possible or probable MTS compared to those without. Subgroup comparisons may be underpowered in light of limited sample sizes (*n* of 6 to 10); however, these null findings are consistent with similar investigations in the literature. St-Laurent and colleagues (2014) found no differences in recall between their participants with left and right TLE and those who were pre- or postsurgical. Similarly, in the autobiographical-memory literature, seizure laterality and presence of MTS are not typically shown to influence recall (Herfurth et al., 2010; McAndrews & Cohn, 2012; Viskontas et al., 2000). As well, in a functional neuroimaging study of hippocampal activation during viewing of the same movie stimulus used here, researchers found no difference in the left and right hippocampus on the effects of interest (the presence and salience of event boundaries in the movie; Ben-Yakov & Henson, 2018). Therefore, the task does not appear to elicit contributions from one hemisphere or the other preferentially, and evidence of a structural lesion on MRI may not disadvantage performance. This latter observation has implications for non-

lesional or “MRI-negative” epilepsies in which the epileptogenic region must be inferred from other investigations beyond structural neuroimaging, including neuropsychology.

To investigate the relative value of the movie-memory measures, each was entered into a regression analysis to predict group membership (TLE vs. HC_{TLE}). Recall of internal details and sensitivity on familiarity judgements were significant predictors, but accuracy on the comprehension task was not found to significantly predict group above and beyond the contributions of the other measures, and it was eliminated from the model. The value of the comprehension task may be as a screening or performance-validity measure to assess, at a basic level, whether the participants were able to pay attention to the movie, follow the plot, and/or comply with task demands. However, this task does not appear to be necessary to detect memory disturbances in TLE, and could, therefore, be eliminated from the protocol to save time. A limitation of this analysis is that the subsections of the movie-memory task were always administered in the same order, since the amount of information provided by the examiner progressively increased with each task. This order may have rendered the comprehension task redundant, if the participant’s memory had been adequately tapped by this point in testing.

3.3.4 Conclusion & future directions

This study provides preliminary support for the use of movie-based memory measures to assess medial temporal-lobe dysfunction in refractory TLE. The question remains as to whether measures of movie memory (or other complex, perceptually rich, naturalistic stimuli) could contribute valuable information to presurgical investigations in TLE that is not currently being captured using simple, unimodal stimuli. For example, although unimodal stimuli can assist with lateralization of the seizure focus (in many cases, but see discussion of material-specific memory above), brain changes in refractory TLE often extend well beyond the seizure focus (Concha, Beaulieu, Collins, & Gross, 2009; Diniz et al., 2011; Liu, Chen, Beaulieu, & Gross, 2014). Tasks that elicit bitemporal involvement may prove useful in measuring broader temporal lobe functioning, which may further be investigated as a predictor of postoperative decline. Another potential application of this

task may be to pair it with a movie-driven functional neuroimaging paradigm. There is a growing trend in the cognitive-neuroscience literature to use movies and other naturalistic stimuli in functional neuroimaging studies, in order to better understand the naturally behaving brain (Hasson & Malach, 2008; Maguire, 2012). An integrated functional magnetic resonance imaging/neuropsychology paradigm would provide an opportunity to investigate direct associations between brain and behaviour, and may ultimately serve as a non-invasive and cost-effective alternative (or complement) to traditional investigation techniques, particularly in more complex surgical cases.

Chapter 4

4 Memory for temporal context in temporal lobe epilepsy

Episodic memory encompasses not only the 'who' and 'what,' but also the 'where' and 'when' of an event. The hippocampus has long been implicated in associative binding of events and their spatiotemporal context (Davachi, 2006; Eichenbaum, 2000; Howard et al., 2014; Mayes, Montaldi, & Migo, 2007; Squire, 1992; Winocur & Moscovitch, 2011), but much of this research has been focused on spatial context (e.g., Bird & Burgess, 2008; Burgess, Becker, King, & O'Keefe, 2001; Chun & Phelps, 1999; Newcombe, Ratcliff, & Damasio, 1987; O'Keefe & Nadel, 1978; Ross & Slotnick, 2008). There is now substantial evidence linking temporal aspects of memory to the hippocampus (Davachi & DuBrow, 2015; Ekstrom & Ranganath, 2018; Howard & Eichenbaum, 2013; Ranganath & Hsieh, 2016), though the potential clinical application of these findings in the neuropsychological assessment of individuals with temporal-lobe lesions remains largely unexplored. Here we consider the potential utility of lab/clinic-based tests of temporal memory in the assessment of individuals with medically refractory temporal lobe epilepsy (TLE).

TLE, and in particular mesial TLE, is the most common type of focal epilepsy and the most likely to be medically refractory (Semah et al., 1998; Wass, Rajala, Hughes, & Sharbrough, 1996). When antiepileptic drugs fail to adequately control seizures, a surgical resection of the epileptogenic tissue is often considered. The main goals of presurgical neuropsychological assessment in individuals with epilepsy are to assist with localization/lateralization of the seizure focus and to predict postoperative changes to cognition (McAndrews & Cohn, 2012). When a temporal-lobe focus is suspected, domains of language and memory are tested extensively to investigate whether seizure activity or a seizure-generating lesion in the temporal lobe has affected these functions, or whether preserved functioning may point to unaffected areas (areas that could potentially bolster functioning postsurgically). The neuropsychological assessment capitalizes on the expected relationship of brain structures and their functions to assist with localization and

lateralization of seizure foci within the temporal lobe. Most notably, verbal and visual episodic memory deficits are thought to implicate left- (or language-dominant) and right-hemisphere involvement, respectively (Gleissner et al., 1998; Glosser et al., 1995; Helmstaedter et al., 1991; Hermann et al., 1997; Kim et al., 2003; Milner, 1972; Sass et al., 1995). Traditional neuropsychological tests use simple stimuli that do not allow for the measurement of temporal aspects of memory.

With memory testing weighted towards stimuli that are primarily encodable in either the verbal or visual modality, memory for rich, naturalistic stimuli is not typically assessed. Tests of naturalistic stimuli would be expected to reflect everyday memory abilities more closely than tests of simpler stimuli, and may capture aspects of cognition that are not currently measured, potentially enhancing the overall sensitivity of the assessment to detect cognitive abnormalities. The focus of the present study is to investigate whether measures of temporal memory for an audiovisual movie stimulus are sensitive to memory deficits in TLE that may not be captured by standardized neuropsychological testing.

Building on foundational animal research demonstrating a critical role for the hippocampus in temporal aspects of memory (Chiba, Kesner, & Reynolds, 1994; Fortin, Agster, & Eichenbaum, 2002; Kesner, Gilbert, & Barua, 2002; MacDonald, Lepage, Eden, & Eichenbaum, 2011; Manns, Howard, & Eichenbaum, 2007; Pastalkova, Itskov, Amarasingham, & Buzsaki, 2008), neuroimaging research has revealed a link between human temporal-lobe function and temporal memory, including memory for temporal context, sequence, duration, and interval length. Most animal and human research has focused on sequence memory (Davachi & DuBrow, 2015; Eichenbaum, 2014). In general, greater medial temporal-lobe activation is associated with better sequence memory (Dudukovic & Wagner, 2007; Jacques, Rubin, Labar, & Cabeza, 2008; Jenkins & Ranganath, 2010; Konishi, Asari, Jimura, Chikazoe, & Miyashita, 2006; Lehn et al., 2009; Ranganath & Hsieh, 2016; Tubridy & Davachi, 2011). The hippocampus plays an important role in predicting successive items in a sequence (Paz et al., 2010; Turk-Browne, Scholl, Johnson, & Chun, 2010), binding consecutive items over temporal gaps

(Hales & Brewer, 2010, 2011; Qin et al., 2007; Staresina & Davachi, 2009), and linking items with their temporal position in a sequence (Hsieh, Gruber, Jenkins, & Ranganath, 2014). In recent years, there has also been a growing interest in the resolution of temporal memory with respect to duration and interval estimation. In humans, the hippocampus and other medial temporal areas are sensitive to duration and interval information on the order of seconds (Barnett, O'Neil, Watson, & Lee, 2014; Thavabalasingam, O'Neil, Tay, Nestor, & Lee, 2019) and minutes (Montchal et al., 2019). Together, these studies demonstrate the role of medial temporal structures in the encoding of the temporal context of episodic memories.

Event boundaries are changes in internal or external contextual states that are perceived at the time of encoding and may serve to segment temporally extended memories into episodes and subepisodes (Clewett et al., 2019; Kurby & Zacks, 2008; Zacks et al., 2007). Behaviourally, event boundaries influence the temporal cohesion of episodic memory, such that details or items that appear in the same context are more likely to be remembered in the correct sequence (DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011; Heusser et al., 2018; Zwaan, 1996) and are more likely to be judged retrospectively as having occurred closer together in time (Ezzyat & Davachi, 2014; Faber & Gennari, 2015), compared to details that cross event boundaries. There is substantial evidence linking hippocampal activity to the perception of event boundaries, suggesting that the hippocampus plays an important role in proactively integrating information within a shared context and separating information across context shifts (Clewett et al., 2019). Hippocampal activity has been shown to peak at the offset of discrete events (e.g., brief film clips; Ben-Yakov & Dudai, 2011; Ben-Yakov et al., 2013) and following event boundaries in continuous audiovisual stimuli (Baldassano et al., 2017; Ezzyat & Davachi, 2011). Ben-Yakov and colleagues (2018) demonstrated that hippocampal activity was not only responsive to the presence of event boundaries (as annotated by multiple observers), it was uniquely modulated by boundary strength (reflected by the extent of agreement across observer ratings). Through reverse inference, they further showed that peak hippocampal activity coincided with subjective annotations of event boundaries. Just as

fluctuating hippocampal activity is associated with temporal separation, stable activity has been linked to increased temporal integration (Deuker, Bellmund, Navarro Schröder, & Doeller, 2016; Dubrow & Davachi, 2014; Ezzyat & Davachi, 2014).

Studies of clinical samples with temporal-lobe lesions further demonstrate the reliance of temporal memory on temporal-lobe integrity. On tests of time estimation, individuals with bilateral medial temporal-lobe damage may underestimate the duration of events, if the duration exceeds the constraints of immediate memory (Perbal, Pouthas, & Van Der Linden, 2000; Williams, Medwedeff, & Haban, 1989). In one study, individuals with unilateral lesions showed intact *production* of duration estimates (producing an event duration given the duration in seconds), but right-sided temporal lesions were associated with underestimates of duration *reproductions* (reproducing an event duration immediately following the target event; Perbal, Ehrlé, Samson, Baulac, & Pouthas, 2001). Studies of relational or context-dependent aspects of episodic memory have identified deficits in sequence memory among individuals with medial temporal-lobe lesions (Konkel, 2008; Spiers, Burgess, Hartley, Vargha-Khadem, & O'Keefe, 2001). Furthermore, in a study of individuals with mild dementia, deficits in sequence memory coincided with aberrant judgements of event boundaries (Zacks et al., 2006).

Several studies have collected temporal-memory measures to better understand autobiographical memory deficits among individuals with temporal-lobe lesions. St-Laurent and colleagues (2011) asked participants with unilateral TLE (pre- and post-anterior temporal lobectomy) to recall specific autobiographical events that occurred one to ten years prior to their visit, and their recall was scored for temporal resolution (based on the number of details reported at different levels of precision), ordering (based on number of apparent sequencing errors), and coherence (based on subjective rating). Compared to control participants, those with TLE produced fewer temporally fine-grained details – thought to reflect a general lack of vividness in recall – but intact sequence memory. To impose additional experimental control, other investigations have used experience-sampling methods (Thaiss & Petrides, 2008) or staged real-world events

(Dede, Frascino, Wixted, & Squire, 2016) to test the veracity of autobiographical recall in clinical samples. Thaiss and Petrides (2008) showed that unlike individuals with frontal-lobe lesions, those with temporal-lobe lesions could not improve upon their event recall when instructed to use a temporal organization strategy, and individuals with left temporal-lobe lesions had the greatest difficulty among individuals with temporal- and frontal-lobe lesions in recalling the temporal sequence of recent autobiographical events. Dede and colleagues (2016) showed that individuals with hippocampal damage were significantly impaired in sequencing recent autobiographical events, and this deficit could not be attributed to weak overall memory alone. Thus, both of these studies test memory against a known sequence of events, but unlike St.-Laurent and colleagues, they miss an opportunity to investigate the resolution of temporal memory.

Naturalistic stimuli, such as movies or virtual-reality environments, are widely used in studies of temporal memory (Baldassano et al., 2017; Ben-Yakov & Henson, 2018; Brunec, Ozubko, Barense, & Moscovitch, 2017; Dede et al., 2016; Deuker et al., 2016; Lehn et al., 2009; Montchal et al., 2019; Paz et al., 2010). These perceptually rich and engaging stimuli approximate the cognitive demands of real-world autobiographical events (Hasson et al., 2010) in ways that simpler, artificial stimuli do not (Gilboa, 2004; McDermott, Szpunar, & Christ, 2009). Tests of autobiographical memory retrieval are highly sensitive to hippocampal dysfunction (Addis et al., 2007; Herfurth et al., 2010; St-Laurent et al., 2009). Unlike autobiographical memory paradigms, however, the accuracy of memories in naturalistic stimulation paradigms can be verified instead of presumed. For example, Thaiss & Petrides (2008) found that individuals with left temporal-lobe lesions produced an elevated number of plausible intrusion errors, details that could not be verified or discounted based on the record of events sampled. Using a single naturalistic stimulus across participants also allows for control of event characteristics such as content, complexity, emotionality, and personal relevance (St-Laurent, Moscovitch, et al., 2014). Furthermore, audiovisual movies represent temporally extended events, which, in itself, offers several advantages. Different brain regions may be tuned to time windows of varying lengths; therefore, longer stimuli are needed to fully

understand the contributions of areas that are influenced by the accumulation of information over longer time windows (Hasson, Yang, et al., 2008). Extended narratives also contain numerous perceptual and narrative contextual shifts, with scene changes thought to represent salient event boundaries that can be reliably identified by independent viewers (Baldassano et al., 2017; Ben-Yakov & Henson, 2018). Thus, naturalistic stimuli provide much of the rich, multimodal stimulation of autobiographical events, with the benefit of greater experimental control.

The goal of the present study was to investigate the sensitivity of temporal-memory measures, derived from a movie-viewing paradigm, to temporal-lobe dysfunction in refractory TLE. Specifically, measures of temporal resolution and temporal sequencing were extracted from various question types designed to probe memory for a brief movie stimulus. Data were collected in two phases. First, in a neurologically healthy control sample of young adults (HC_{YA}), we sought to investigate the reliability and construct validity of our temporal-memory measures. Next, we recruited a sample of individuals with refractory TLE and demographically matched healthy controls (HC_{TLE}) to reproduce these reliability and validity investigations in the clinical sample and to investigate the sensitivity of these measures in distinguishing TLE from control participants. Based on these investigations, we hope to better understand the value of temporal-memory measures in the presurgical neuropsychological work-up of individuals with TLE to enhance detection of temporal-lobe dysfunction.

4.1 Methods

4.1.1 Healthy validation

4.1.1.1 Participants (HC_{YA})

Twenty-four participants with no history of neurological or psychiatric illness were recruited for this phase of the study. All participants were native English speakers (defined as having learned English before age 5), with no hearing difficulties (based on self-report), and normal or corrected-to-normal vision. This sample was recruited

predominantly from Western University's undergraduate population, and as a result, the group is relatively young and well-educated. See Table 7 for more demographic information.

Table 7. Participant characteristics and cognitive performance.

	HC_{YA}	HC_{TLE}	TLE
n	24	23	16
Demographic and clinical characteristics			
Sex (F:M)	14:10	13:10	8:8
Age (*M* ± *SD*)	23.17 ± 3.24	35.65 ± 15.22	33.88 ± 11.85
Years of Education (*M* ± *SD*)	15.63 ± 2.32	13.64 ± 1.73	12.56 ± 2.28
Handedness (R:L)	24:0	18:5	16:0
Seizure Lateralization			6 L; 9 R; 1 BL
Years since onset (*M* ± *SD*)			19.07 ± 17.66
Standardized test performance: Median (Min. - Max.)			
SAM Episodic	97 (78 - 128)		
Vocabulary (scaled)		13 (9 - 19)	9 (3 - 12)
Matrix Reasoning (scaled)	12 (8 - 16)	11 (6 - 16)	10 (4 - 12)
Logical Memory (scaled)	12 (7 - 15)		7 (1 - 15)
CVLT			8 (0 - 15)
RAVLT	13 (8 - 15)		
RVDLT		12 (5 - 15)	11 (4 - 15)
Names		21 (14 - 23)	19 (14 - 23)
RCFT			13 (6.5 - 23.5)
Doors		22 (13 - 24)	19 (9 - 22)
CALT	26 (12 - 68)	32 (14 - 68)	41 (17 - 68)

	HC_{YA}	HC_{TLE}	TLE
Temporal memory performance: Median (Min. - Max.)			
Recall			
Clustered Details	21 (9 - 60)	29 (6 - 66)	18.5 (1 - 42)
Indefinite & Higher Order Details	49 (30 - 67)	56 (28 - 82)	35 (22 - 93)
Sequencing Errors	3 (0 - 5)	4 (2 - 8)	4 (0 - 5)
Recall Composite	72 (44 - 129)	80 (45 - 152)	57 (23 - 140)
Timeline			
Position estimation	60 (36 - 142)	63 (44 - 99)	79 (47 - 132)
Interval estimation	45 (28 - 74)	48 (33 - 76)	71 (39 - 108)
Correctly Ordered Trials	.86 (.64 - .96)	.82 (.61 - .93)	.75 (.54 - .89)
Timeline Composite	.21 (.10 - .37)	.19 (.11 - .29)	.13 (.08 - .28)
Recall & Timeline Composite	.91 (.42 - 1.81)	.90 (.38 - 1.70)	.51 (.19 - 1.15)

Note. All standardized measures are reported as raw scores, except the age-scaled scores reported for Vocabulary, Matrix Reasoning, and Logical Memory. The TLE sample received the WASI-II version of Vocabulary and Matrix Reasoning and the WMS-III version of Logical Memory. The healthy control groups received the WAIS-IV and WMS-IV versions of these tests. HC_{YA} = healthy control participants – young adults; HC_{TLE} = healthy control participants – matched to epilepsy sample; TLE = temporal lobe epilepsy sample; F = female; M = male; R = right; L = left; BL = bilateral; SAM = Survey of Autobiographical Memory; CVLT = California Verbal Learning Test; RAVLT = Rey Auditory Verbal Learning Test; RVDLT = Rey Visual Design Learning Test; RCFT = Rey Complex Figure Test; CALT = Conditional Associative Learning Test.

4.1.1.2 Procedure

Participants were asked to undergo functional magnetic resonance imaging, the results of which are not reported here. During the last imaging run, they were shown the movie clip, and upon exiting the scanner, they were asked to complete a memory test for the movie. Prior to viewing the movie, they were instructed to pay attention and follow the plot of the movie as they would if they were watching any movie or television show. They were not warned that after the imaging session, there would be a test of memory for the movie. The temporal-memory measures were derived from two parts of the movie memory test – *free and probed recall* and *timeline judgements*. Participants were also asked to complete a brief battery of standardized cognitive tests.

4.1.1.3 Movie stimulus

The movie stimulus was an eight-minute long, black-and-white, audio-visual clip. The clip was edited from a 20-minute television episode entitled “Alfred Hitchcock Presents: Bang! You’re Dead.” The clip depicts a young boy who finds his uncle’s revolver. He goes around town and plays “cowboy” with the real gun, believing it to be a toy, unaware of the danger. The episode was originally broadcast in 1961 and all participants (in the healthy- and clinical-validation studies) denied having ever seen it before the scanning session. The identical edited clip has been used in other neuroimaging studies (Ben-Yakov & Henson, 2018; K. L. Campbell et al., 2015; Hasson et al., 2010; Naci et al., 2014; Taylor et al., 2017). The suspenseful plot, as well as its novelty, was expected to promote engagement and interest (Naci et al., 2014).

4.1.1.4 Free and probed recall

4.1.1.4.1 Administration

Free and probed recall followed a structured-interview format and was audio-recorded for later scoring. The interview, based on the staged probing procedure of the Autobiographical Interview (AI; Levine et al., 2002), was adapted to evaluate memory for the movie instead of memory for autobiographical events. In *free recall*, participants were instructed to provide as much detail as they could about the movie. In probed recall, participants were prompted to provide any more information (*general probe*) and were then asked a series of questions pertaining to various aspects of the movie, such as characters, settings, and perceptual details (*specific probes*). All specific probes were administered; however, responses that were already provided during an earlier recall condition were ignored during scoring as if the probe was not administered.

4.1.1.4.2 Scoring

All free and probed recall measures are referred to as “recall” scores for brevity. Audio recordings were transcribed verbatim. A sample scored transcript is displayed in Figure 5. Transcripts were segmented into individual details (i.e., single units of information),

modeled after the text-segmentation procedure described for the AI. Also based on the AI, details were classified as *internal* – directly relevant to the main event (in this case, the content and experience of watching the movie) – and *external* – not directly relevant to the main event, semantic information not specific to the main event, or repetition of internal details. Internal details were further classified as *correct* or *incorrect*, where incorrect details were those that unambiguously opposed information available in the movie. Only correct internal details were considered in the evaluation of temporal aspects of memory.

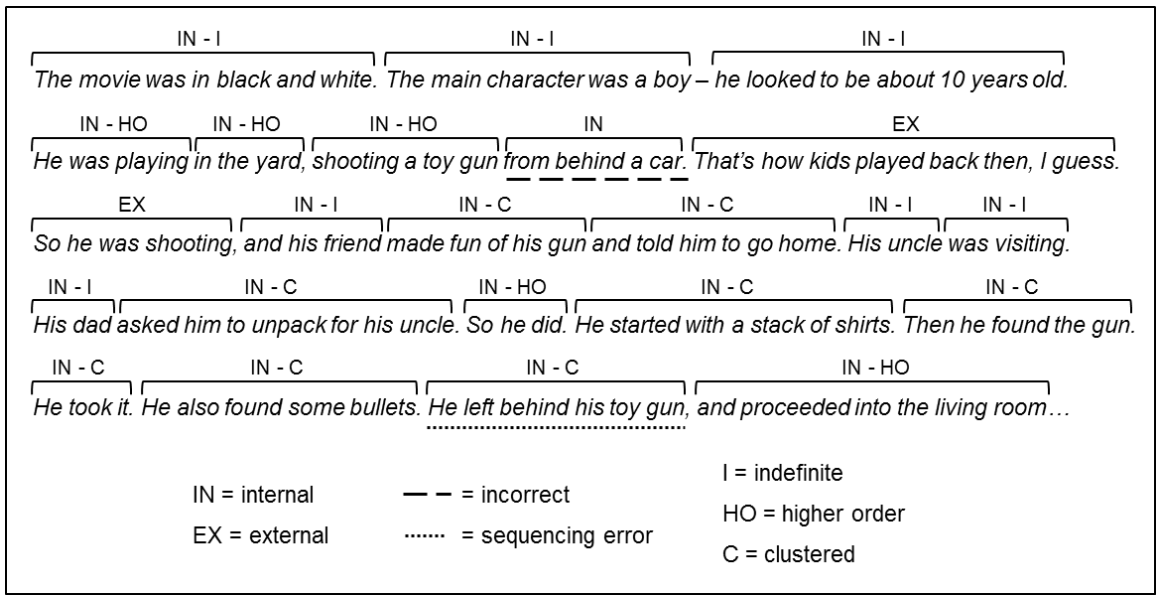


Figure 5. A sample transcript of a partial recall response, showing segmentation and classification of details types and errors.

Scoring of temporal-memory measures was adapted from the procedure devised by St-Laurent and colleagues (2011) for scoring autobiographical-memory narratives that were collected based on the AI (see Appendix C). Briefly, to investigate the temporal resolution of memory, St-Laurent and colleagues (2011) classified internal details as temporally indefinite (pertaining to the entire length of the main event), temporally precise – higher order (pertaining to subepisodes of the main event), and temporally precise – clustered (pertaining to shorter-duration, specific actions). Applying this

procedure to the movie recall, *indefinite* details were those that pertained to the entirety of the movie, *higher order* details pertained to subevents that spanned one or more scenes but retained a clear beginning and/or end, and *clustered* details pertained to specific subevents that unraveled on the order of seconds to minutes. Each detail type was tallied across all three of the recall conditions (free recall, general probe, specific probes).

To evaluate accuracy of temporal ordering, St-Laurent and colleagues (2011) tallied the breaks in chronology of the reported events in the autobiographical-memory narrative. We, too, tallied *sequencing errors* in this way. Unlike the autobiographical-memory paradigm, however, we were able to judge chronology against the unfolding of events in the movie, rather than inferring it from context. A sequencing error was counted whether or not it was self-corrected (e.g., when the participant stated, “Before that…,” “I forgot to mention…,” and the like). The number of sequencing errors was tallied continuously across the free-recall and general-probe conditions. Therefore, if new information was provided after the general probe that did not follow chronologically from the last higher order or clustered detail produced during free recall, a sequencing error was counted. The specific-probe condition was not coded for sequencing errors, as the sequencing of events in this condition was guided externally by the interview and not by an internally generated chronology.

As the recall task used a subjective scoring system, inter-rater reliability was assessed across two independent raters. Reliability was investigated for the three detail types (indefinite, higher order, clustered) and sequencing errors. Intraclass correlation coefficients were calculated based on a single-rating, absolute-agreement, two-way mixed effects model, assuming the intended use of these measures as being rated by a single clinician/researcher (McGraw & Wong, 1996). Raters applied the scoring guidelines to each of the 24 transcripts of the healthy validation study, including segmenting and classifying details. After a discussion of scoring discrepancies, the guidelines were revised and reapplied to the same 24 transcripts. The scores of a single rater were used for analysis.

Two sets of intraclass correlation coefficients for the healthy-validation study, before and after revising the scoring guidelines, are displayed in Table 8. It was difficult to establish high inter-rater reliability for indefinite details (intraclass correlation of 0.52 even after revising scoring guidelines). Whereas the theoretical distinction of temporally indefinite and higher order details was readily apparent for ratings of autobiographical memories, the distinction between these variables based on movie-memory performance may be more subtle. That is, indefinite details for the movie (details pertaining to the entirety of the movie) still represent memories of a brief, circumscribed period of time. Indeed, these variables were highly correlated in the healthy-validation sample, $r = 0.62$, $p = 0.001$. Therefore, we thought it prudent to combine indefinite and higher order details to provide a more reliable and theoretically motivated measure of this temporally courser-grained level of memory. This combined measure was used in subsequent analyses. Intraclass correlations for the healthy-validation study following revision of the scoring guidelines demonstrated "fair" to "excellent" reliability (Cicchetti, 1994). Lower coefficients for sequencing errors may be explained by the low base rate of sequencing errors (*M*: 2.79, *SD*: 1.50), making absolute agreement across raters more difficult to achieve.

Table 8. Inter-rater reliability estimates of recall scores based on intraclass correlations.

	Healthy Validation Before revising scoring guidelines	Healthy Validation After revising scoring guidelines	Clinical Validation
Clustered details	.88	.86	.71
Higher order details	.81	.80	
Indefinite details	.45	.52	
Indefinite & higher order details	.69	.74	.75
Sequencing errors	.47	.59	.65

Note. Intraclass correlation coefficients were calculated based on a single-rating, absolute-agreement, two-way mixed-effects model.

4.1.1.5 Timeline judgements

4.1.1.5.1 Administration

The timeline-judgement task was presented via the Psychophysics toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) in MATLAB 2014b (Mathworks, Natick, MA, USA). For each item, participants were shown two still frames from the movie side by side, labeled "1" and "2," as well as a timeline beneath them (see Figure 6A). Participants were instructed to use a computer mouse to click on the timeline where they recalled each still frame to have taken place in the course of the movie. An "X" appeared at the site of each mouse click. Twenty-eight items were presented in total. To screen for motor difficulties, three calibration trials were included, instructing the participant to click at several points along the timeline that were indicated by a red arrow.

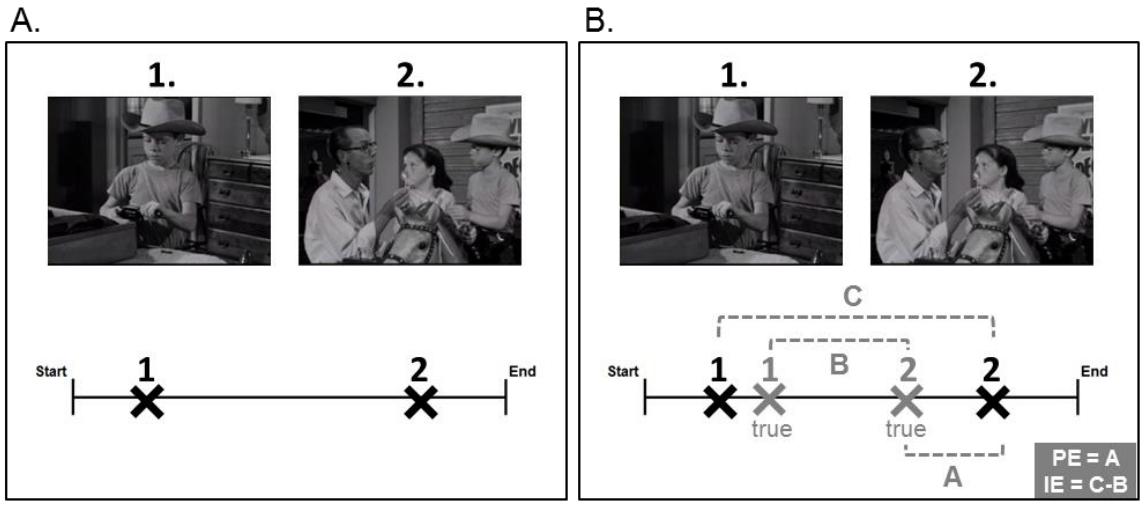


Figure 6. Timeline judgements administration and scoring. *A* depicts a single timeline-judgement item after the participant has provided a response for both still frames ('X's appear where the participant has clicked). *B* depicts the measures used to compute timeline-judgement scores which are averaged across items. PE = position estimation error; IE = interval estimation error.

Selection of the 28 items (composed of 46 still frames) was based on two principles. First, a set of 40 items were piloted in an independent sample of 30 healthy young adults (age $M \pm SD = 18.63 \pm 0.96$; 21 female). Items/still frames that elicited variable performance were favoured over items that demonstrated floor or ceiling effects. Second, items were sampled based on three parameters to comprehensively assess movie memory: (1) the serial position of information in the course of the movie (10-17 still frames sampled from each quarter of the movie), (2) the interval length between still frames in a given item (four items sampled at each of 60, 120, and 240 s intervals and 8 items sampled at each of 15 and 30 s intervals), and (3) whether the two still frames of a given item occurred within the same scene or across different scenes (i.e., across gross event boundaries). The latter distinction was only considered for shorter interval lengths since within-scene pairs could not be sampled at longer intervals; of the 16 items with shorter interval lengths (15 and 30 s), eight were sampled within and eight were sampled across scenes.

4.1.1.5.2 Scoring

See Figure 6B for a graphic representation of scoring. Three main outcome measures were computed: correct ordering, position estimation error, and interval estimation error. *Correct ordering* was measured as the total number of items in which the two still frames were ordered correctly. Unless otherwise indicated, correct ordering was expressed as the proportion of correctly ordered trials over the total number of trials. Displacement in the horizontal direction of each mouse click along the timeline was converted into seconds. *Position estimation error* was defined as the absolute difference between the true position of a still frame in the movie and the position estimated by the participant, averaged across still frames. Since no systematic difference was observed between still frames presented on the left versus those on the right ($p = .208$), their position estimates were pooled. *Interval estimation error* was defined as the absolute difference between the true interval (i.e., the difference in temporal position of the two still frames in the movie) and the interval estimated by the participant, averaged across items. The interval estimated by the participant was also retained to investigate the influence of shared context (whether items

were sampled from the same or different scenes) on interval judgement. Position and interval estimation error were only measured for trials on which the participant correctly ordered the still frames. All participants correctly ordered more than 50% of trials (i.e., chance level responding). These three measures were selected to capture aspects of sequence memory and the temporal resolution of memory. We wished to explore the properties of each, despite their potential overlap.

4.1.1.6 Standardized cognitive tests

Three memory tests were administered. Verbal memory was measured using the delayed recall trial of the *Rey Auditory Verbal Learning Test* (RAVLT; Strauss et al., 2006a), in which participants are asked to learn a 15-item word list over five learning trials and then recall the list again after a 20-minute delay. Contextualized verbal memory was assessed using the delayed recall trial on *Logical Memory* (from the Wechsler Memory Scaled-IV, WMS-IV; Wechsler, 2009), for which participants are asked to learn two short stories and recall them again after a 20- to 30-minute delay. Visual memory was assessed based on the trials-to-criterion measure of the *Conditional Associative Learning Test* (CALT; Petrides, 1985; St-Laurent, McCormick, et al., 2014), a test of spatial associative learning in which participants are asked to learn arbitrary associations between four cards and four spatially dispersed discs, discontinuing after 68 trials or sooner if the criterion of 12 consecutive correct trials is met. The episodic subscale of the *Survey of Autobiographical Memory* (SAM; Palombo et al., 2013) was included as a self-report measure of naturalistic memory abilities. *Matrix Reasoning*, a test of nonverbal problem solving from the Wechsler Adult Intelligence Scale-IV (WAIS-IV; Wechsler, 2008), was included as a "control" measure which was expected to show a weaker relationship with the temporal-memory measures compared to other standardized measures. For two tests – Logical Memory and Matrix Reasoning – two different versions were administered to different study samples (i.e., WMS-IV Logical Memory to HC_{YA} and WMS-III to TLE, WAIS-IV Matrix Reasoning to HC_{YA} and HC_{TLE} and Wechsler Abbreviated Scale of Intelligence-II, WASI-II, to TLE), and so age-scaled scores were used to facilitate

comparison. For all other tests, raw scores were used for analysis. Median scores for each test in the HC_{YA} sample are presented in Table 7.

4.1.2 Clinical validation

4.1.2.1 Participants (TLE and HC_{TLE})

Thirty-nine participants were enrolled in this study phase. Sixteen participants were recruited into the epilepsy sample from the Adult Epilepsy Service at London Health Sciences Centre (University Hospital) in London, Ontario, Canada. All were identified by their neurologist as potential candidates for a temporal-lobe resection to control seizures based on evidence of focal structural and functional abnormalities. None had undergone neurosurgery prior to participation. Among the 16 participants, eight demonstrated evidence of medial temporal sclerosis (MTS) on MRI, three demonstrated equivocal evidence of a temporal-lobe lesion (MTS or focal cortical dysplasia), and five demonstrated no structural abnormalities. Twenty-three participants were recruited into the healthy control sample (HC_{TLE}) from the local community. These participants had no history of neurological or psychiatric illness, were native English speakers (defined as having learning English before age 5), with no hearing difficulties (based on self-report), and normal or corrected-to-normal vision. They were demographically matched to the TLE sample on sex (p = .688), age (p = .698), and education (p = .107). Further demographic and clinical information about these participants can be found in Table 7.

4.1.2.2 Procedure

The study procedure, including administration and scoring of experimental outcome measures, was largely identical to the procedure followed in the healthy-validation study. The few changes to the study procedure are noted below.

4.1.2.3 Free and probed recall

Transcripts were anonymized prior to segmenting and classifying details, so that raters were blind to the experimental group of the participant (TLE or control). To evaluate

inter-rater reliability, the same two raters who rated transcripts for the healthy validation independently rated the transcripts of 20 participants (10 TLE) from the clinical-validation study. Ratings were based on the same scoring guidelines used for the healthy validation (after revision). Intraclass correlations (displayed in Table 8) were "good" to "excellent," with lower reliability for sequencing errors likely reflective of overall low numbers of sequencing errors (M = 3.48, SD = 1.91). The scores of a single rater, who also scored the remaining 20 transcripts from the clinical validation study, were used for analysis.

4.1.2.4 Timeline judgements

Five participants (3 TLE and 2 control) were administered an earlier version of the timeline-judgement task comprised of 20 (instead of 28) items. To improve statistical power, these participants were included in analyses for which overall outcome measures were used (i.e., proportion of correctly ordered trials, average position estimation error across all trials, average interval estimation error across all trials), but were otherwise excluded from analyses related to timeline judgements.

4.1.2.5 Standardized cognitive tests

For 15 of the 16 TLE participants, data from their clinical neuropsychological assessment (part of the presurgical work-up) were available. From this large battery of tests, we obtained measures of verbal memory – delayed recall from *Logical Memory* (WMS-III; Wechsler, 1997) and delayed recall from the *California Verbal Learning Test* (CVLT; Delis et al., 2000) – visual memory – delayed recall on the *Rey Complex Figure Test* (RCFT; Strauss et al., 2006b) – and proxy measures of crystallized/verbal and fluid/visual intelligence – *Vocabulary* and *Matrix Reasoning* from the WASI-II (Wechsler, 2011). In addition, the TLE and HC_{TLE} samples were assessed on delayed recall on the *Rey Visual Design Learning Test* (RVDLT; Spreen & Strauss, 1991), a measure of visual memory for a supraspan list of abstract designs, *Names* and *Doors*, verbal and visual recognition subtests of the Doors & People Test (Baddeley et al., 2006), and the *Conditional Associative Learning Test* (CALT; Petrides, 1985; St-Laurent, McCormick, et al., 2014).

Finally, the HC_{TLE} sample was also administered the *Vocabulary* & *Matrix Reasoning* subscales of the WAIS-IV (Wechsler, 2008). Raw scores were used in analysis, with the exception of Matrix Reasoning, Vocabulary, and Logical Memory, for which age-scaled scores were used to facilitate comparison across the test versions administered to different samples (i.e., WASI-II and WAIS-IV, WMS-III and WMS-IV). Median scores for each test administered to the TLE and HC_{TLE} samples are presented in Table 7.

4.1.3 Analyses

All statistical analyses were performed using IBM SPSS Statistics 25 (IBM, Armonk, NY, USA). To assess the sensitivity of these measures to memory functioning in TLE, a series of Mann-Whitney U Tests were performed comparing TLE and HC_{TLE} samples on temporal and standardized memory measures (under false discovery rate, FDR, correction; Benjamini & Hochberg, 1995). In addition to comparing performance across individual tests, several composite measures were created to magnify sensitivity to group: one for the timeline scores, one for the recall scores, and one combining the timeline and recall scores. The construction of these composite scores is described in more detail below.

Construct validity was investigated in several ways. First, temporal-memory measures were correlated among themselves to investigate the degree to which measures of the same task hang together (reflecting internal consistency) and the degree to which measures of the different tasks covary (reflecting the degree of similarity of target constructs). Second, bivariate correlations among temporal- and standardized-memory measures were used to determine the extent to which the temporal-memory tasks assess similar constructs to standardized cognitive tests. Third, we can capitalize on previous studies of temporal memory to investigate whether our measures respond in expected ways to certain manipulations that have previously been shown to influence temporal memory. Specifically, the influence of shared context/event boundaries on later retrieval of temporal information can be investigated in the timeline-judgement task. We also tested whether performance was influenced by the sampling parameters of the timeline

items, including the quarter of the clip from which still frames were sampled and the true interval at which the two still-frames that make up one item were sampled.

Finally, binary logistic regression was used to predict group membership (TLE vs. HC_{TLE}) from standardized- and temporal-memory measures. Using a hierarchical model, we can investigate the individual variability explained by temporal-memory measures above and beyond standardized tests. Using group-classification statistics, we can also quantify the sensitivity and specificity of these measures.

4.2 Results

4.2.1 Group comparisons

Group comparisons on standardized and temporal-memory measures are shown in Table 9. Among the standardized measures administered to both the TLE and HC_{TLE} groups, only performance on the Doors test was significantly lower in the epilepsy sample than in the matched controls ($U = 60.0$, $p < .001$, $\eta^2 = .31$). Among the temporal-memory measures, all measures derived from the timeline and recall tasks, with the exception of recall sequencing errors, were significantly impaired in the epilepsy sample.

Table 9. Group comparisons across standardized and temporal-memory measures.

	U	*p*	η^2
Standardized			
Doors	60.0	<.001	.31
Names	129.0	.121	.07
RVDLT	123.0	.084	.08
CALT	111.5	.056	.10
Recall			
Sequencing Errors	140.5	.217	.04
Indefinite & Higher Order Details	68.0	.001	.29
Clustered Details	111.0	.037	.11
Recall Composite	**72.5**	**.001**	**.27**
Timeline			
Correct Ordering	85.0	.004	.21
Position Estimation Error	101.0	.017	.15
Interval Estimation Error	88.0	.005	.20
Timeline Composite	**63.0**	**<.001**	**.31**
Recall & Timeline Composite	**66.0**	**<.001**	**.30**

Note. As some measures violated parametric assumptions, the Mann-Whitney U Test was applied to test all group differences for ease of comparison. Composite measures are shown in **bold.** RVDLT = Rey Visual Design Learning Test; CALT = Conditional Associative Learning Test.

Composite measures were computed for timeline measures, recall measures, and their combination, by summing measures that *increase* with better performance and dividing by the sum of measures that *decrease* with better performance. The timeline composite represented the total number of correctly ordered trials, divided by the sum of position and interval estimation error. (Note that the number of correctly ordered trials and not the proportion was used, in order to match the order of magnitude of the other timeline measures.) The recall composite represented the sum of the number of indefinite, higher order, and clustered details (in other words, the total number of details recalled). Sequencing errors were excluded as they were not significantly affected by group, whether tested as a raw number ($p = .217$) or divided by the total number of details recalled (to control for their positive association; $p = .767$). The timeline and recall composite represented the sum of timeline correctly ordered trials, recall of indefinite &

higher order details, and recall of clustered details, divided by the sum of timeline position and timeline interval estimation error. Based on magnitude of effect size (η^2) among the temporal memory composite measures, 27 to 31% of variability in task performance could be accounted for by group differences.

To investigate the influence of seizure laterality on temporal-memory task performance, we separated the TLE group into those with left-hemisphere lateralized foci (left TLE; *n* = 6) and right-hemisphere lateralized foci (right TLE; *n* = 9). No significant differences were detected between the right and left TLE groups on any of the temporal-memory measures. Comparing these subgroups to the control group, we found a similar pattern of group differences to those observed in the entire TLE sample (see Table 10). That is, several individual temporal-memory measures and all composite measures significantly distinguished both TLE subgroups from the control group, despite relatively small sample sizes.

Table 10. Group comparisons between clinical subgroups and controls.

	Left TLE (*n* = 6)			Right TLE (*n* = 9)		
	U	*p*	η^2	*U*	*p*	η^2
Recall						
Sequencing Errors	51.5	.356	.03	89.0	.564	.01
Indefinite & Higher Order Details	17.5	.003	.28	50.5	.024	.16
Clustered Details	33.0	.054	.13	78.0	.301	.04
Recall Composite	**19.0**	**.005**	**.26**	**53.5**	**.034**	**.14**
Timeline						
Correct Ordering	23.5	.011	.22	59.5	.064	.11
Position Estimation Error	36.0	.080	.11	62.0	.086	.10
Interval Estimation Error	20.0	.006	.25	67.0	.133	.08
Timeline Composite	**19.0**	**.005**	**.26**	**43.0**	**.010**	**.21**
Recall & Timeline Composite	**16.0**	**.003**	**.29**	**50.0**	**.024**	**.16**

Note. Group differences were evaluated by the Mann-Whitney U Test. Composite measures are shown in **bold**. TLE = temporal lobe epilepsy.

4.2.2 Correlations among recall and timeline measures

Temporal-memory measures were correlated within and across the recall and timeline tasks, separately for healthy and epilepsy samples. In order to pool data across the two healthy samples (HC_{YA} and HC_{TLE}), we first conducted correlations in each healthy sample independently and then tested them for significant group differences. A 95% confidence interval (CI) for the difference between the correlations was computed, where an interval that includes 0 suggests a nonsignificant difference (Zou, 2007). Only one correlation differed significantly between the two HC groups – that between timeline position estimation error and recall of indefinite & higher order details (95% CI: [0.02, 1.06]; more strongly negative in the HC_{TLE} group). Therefore, we can be reasonably confident in pooling these two healthy samples in subsequent correlational analyses.

Examining correlations within and across temporal-memory tasks in this pooled healthy sample (n = 47), three significant correlations survived FDR correction: timeline position estimation error and timeline interval estimation error (r_s = 0.62, p < .001), recall of indefinite & higher order details and recall of clustered details (r_s = 0.48, p < .001), and recall of indefinite & higher order details and recall sequencing errors (r = 0.49, p < .001). This latter correlation may reflect the fact that producing more details during recall gives the participant more opportunity to make sequencing errors; therefore, more sequencing errors may be suggestive of better performance. All three significant correlations were within measures of the same task (timeline or recall); no correlations across temporal-memory tasks survived correction.

Computing correlations in the TLE sample, we found a similar pattern of results as in the healthy samples. The three correlations to survive FDR correction were among the measures of the recall task: indefinite & higher order and clustered details (r_s = 0.70, p = .002), indefinite & higher order details and sequencing errors (r_s = 0.83, p < .001), clustered details and sequencing errors (r_s = 0.78, p < .001). No within-task correlations among timeline measures, or across-task correlations, survived correction. Comparing correlations across healthy and epilepsy samples, we found three correlations that were

significantly stronger in the TLE sample (and shared the same direction of correlation across group): the positive correlation between recall sequencing errors and indefinite & higher order details (95% CI: [0.02, 0.62]), the positive correlation between recall sequencing errors and clustered details (95% CI: [0.08, 0.80]), and the negative correlation between recall sequencing errors and timeline position estimation error (95% CI: [0.02, 0.96]). In summary, the positive relationship between sequencing errors and other recall and timeline measures reflective of better performance was even more pronounced in the TLE sample. See Figure 7 for the relevant scatterplots.

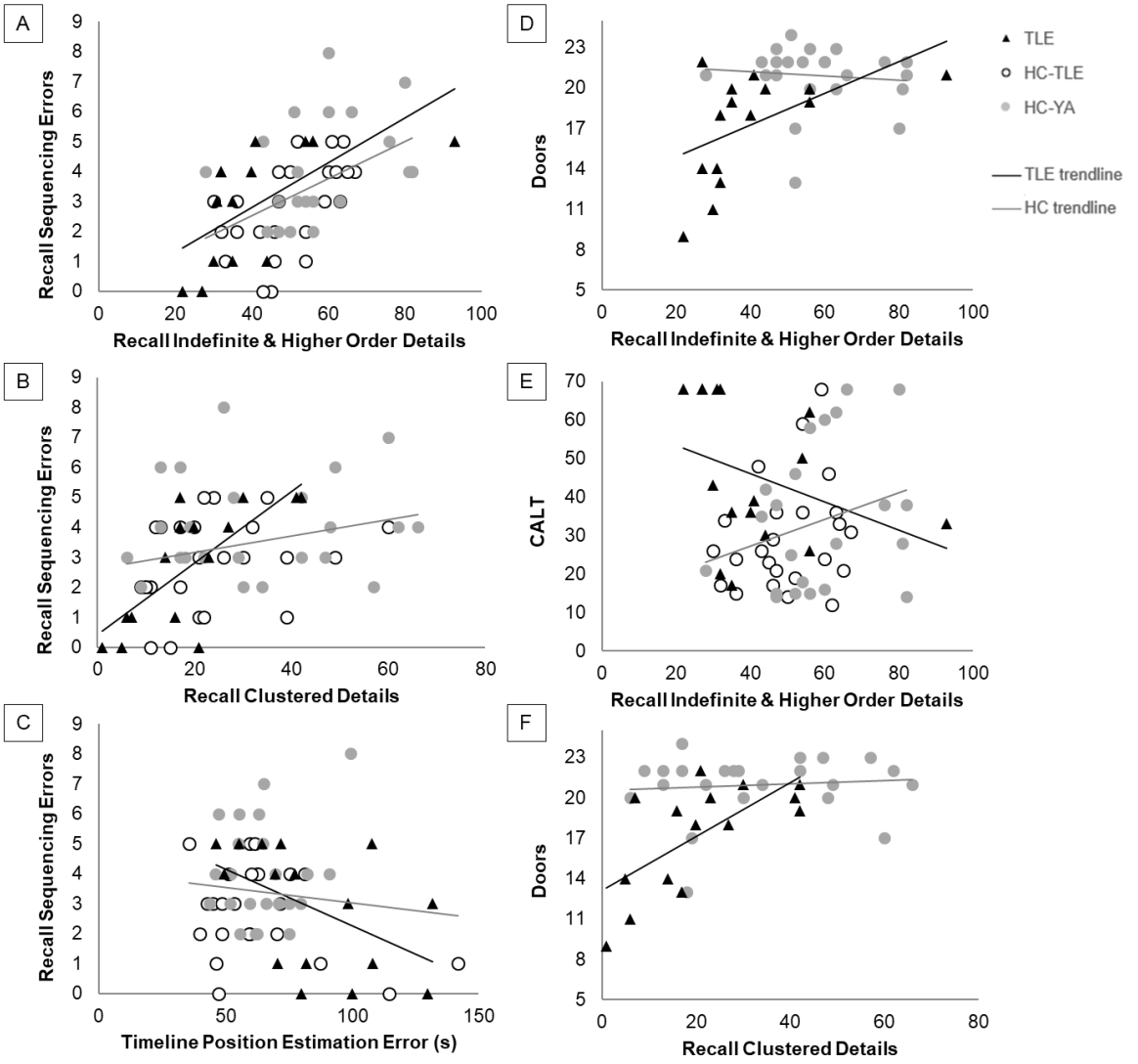


Figure 7. Scatterplots representing the correlations which were significantly different between the TLE and pooled control samples (HC_{TLE} and HC_{YA}). All correlations were stronger in the TLE sample. *A* to *C* depict correlations among the temporal-memory measures. *D* to *F* depict correlations between temporal-memory and standardized measures. TLE = temporal lobe epilepsy sample; HC_{TLE} = healthy control participants – matched to epilepsy sample; HC_{YA} = healthy control participants – young adults; CALT = Conditional Associative Learning Test.

Finally, in an analysis aimed at maximizing sensitivity to associations common across the groups, we pooled all three study samples (N = 63) and re-examined correlations (with FDR correction). The full array of correlations is displayed in Table 11. All three recall measures were highly positively correlated (r_s = 0.43 to 0.55, p < .001). Among timeline measures, position and interval estimation error were highly positively correlated (r_s = 0.66, p < .001) and interval estimation error and correctly ordered trials were moderately negatively correlated (r_s = -0.33, p = .001). Between the timeline and recall measures, four significant correlations emerged that did not survive correction in the subgroups: timeline position estimation error was negatively correlated with indefinite & higher order details (r_s = -0.35, p = .005) and clustered details (r_s = -0.33, p = .008), and timeline correctly ordered trials was positively correlated to these same variables (indefinite & higher order: r_s = 0.41, p = .001; clustered: r_s = 0.30, p = .017).

Table 11. Correlations among temporal-memory measures.

	Recall						Timeline			
	Clustered Details		Indefinite & Higher Order Details		Sequencing Errors		Position Estimation		Interval Estimation	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Indefinite & Higher Order Details	**.54****	**.001**								
Sequencing Errors	**.43****	**<.001**	**.55****	**<.001**						
Position Estimation	**-.33****	**.008**	**-.35****	**.005**	-.20	.116				
Interval Estimation	-.25	.051	-.27*	.034	-.08	.518	**.66****	**<.001**		
Correctly Ordered Trials	**.30****	**.017**	**.41****	**.001**	.19	.134	-.24	.054	**-.33****	**.008**

Note. Correlations are collapsed across the three study samples (N = 63). Bolded correlations are those that survived FDR correction.
*significant at the .05 level
**significant at the .01 level

4.2.3 Correlations across temporal-memory and standardized measures

Just as before, we compared correlations in each sample of healthy participants to justify pooling them. The healthy samples shared two standardized measures – Matrix Reasoning and the Conditional Associative Learning Test. Comparing the correlations of these tests with temporal-memory measures, we found no significant differences across the two control samples. Applying FDR correction to the full set of correlations available across the healthy samples, no correlations remained significant.

Next, we conducted the same correlations in the TLE sample. Among correlations with recall measures, several significant correlations emerged with Logical Memory, Doors, CALT, and CVLT; however, none survived correction. Similarly, timeline measures correlated with Logical Memory, RCFT, CVLT, RVDLT, and Doors, yet none survived correction. Statistically comparing correlations across the TLE and pooled healthy samples, we found no differences in the strength of correlations among timeline and standardized measures. We did, however, find 3 correlations with recall measures that were significantly stronger in the TLE group: indefinite & higher order details was significantly more correlated with Doors (95% CI: [0.08, 1.19]) and with CALT (95% CI: [0.24, 1.17]), and clustered details was significantly more correlated with Doors (95% CI: [0.05, 1.10]). In other words, the association between number of details recalled and standardized measures of visual-spatial memory was stronger in the TLE group. See Figure 7 for the relevant scatterplots.

When we combined all three study samples to improve statistical power, we found no significant correlations across recall and standardized measures that survive FDR correction. However, all timeline measures demonstrated moderate to strong correlations with standardized measures. Position estimation error significantly correlated with Logical Memory (r_s = -0.60, p < .001), Vocabulary (r_s = -0.40, p = .012), and CVLT (r_s = -0.65, p = .008). Interval estimation error correlated with Logical Memory (r_s = -0.40, p = .012), RVDLT (r_s = -0.45, p = .004), and Doors (r_s = -0.52, p = .001). Correctly ordered

trials correlated with Logical Memory ($r_s = 0.47$, $p = .003$), Doors ($r_s = 0.58$, $p < .001$), Vocabulary ($r_s = 0.39$, $p = .015$), and RCFT ($r_s = 0.69$, $p = .004$). In summary, timeline measures correlated widely with measures of visual memory, verbal memory, and verbal intelligence. The full array of correlations is shown in Table 12.

Table 12. Correlations between temporal-memory measures and standardized cognitive measures.

		Recall						Timeline					
		Clustered Details		Indefinite & Higher Order Details		Sequencing Errors		Position Estimation		Interval Estimation		Correctly Ordered Trials	
	n	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Logical Memory	39	.36*	.023	.44*	.005	.23	.154	**-.60***	**.000**	**-.40***	**.012**	**.47***	**.003**
RCFT	15	.36	.190	.31	.266	.19	.488	-.48	.072	**-.**47	.077	**.69***	**.004**
CVLT	15	.54*	.040	.40	.143	.57*	.028	**-.65***	**.008**	**-.**20	.466	.49	.065
RAVLT	24	.31	.143	.15	.492	.17	.434	-.26	.223	-.19	.377	.14	.527
RVDLT	39	.13	.446	.27	.102	.06	.717	-.21	.191	**-.45***	**.004**	.34*	.036
Names	39	-.04	.810	.07	.686	-.06	.719	-.03	.875	-.15	.359	.09	.592
Doors	38	.38*	.021	.37*	.021	.21	.208	-.37*	.022	**-.52***	**.001**	**.58***	**.000**
CALT	62	-.12	.342	-.04	.766	-.01	.930	.11	.414	.28*	.027	-.14	.269
Vocabulary	38	.34*	.040	.49*	.002	.22	.180	**-.40***	**.012**	-.36*	.029	**.39***	**.015**
Matrix Reasoning	62	-.03	.819	.11	.406	.08	.541	-.12	.342	-.25*	.047	.09	.501
SAM Episodic	24	-.25	.235	-.11	.607	-.20	.341	.36	.082	.37	.079	-.45*	.027

Note. Correlations are collapsed across study samples for tests administered to more than one sample. Bolded correlations are those that survived FDR correction. RCFT = Rey Complex Figure Test; CVLT = California Verbal Learning Test; RAVLT = Rey Auditory Verbal Learning Test; RVDLT = Rey Visual Design Learning Test; CALT = Conditional Associative Learning Test; SAM = Survey of Autobiographical Memory.
*significant at the .05 level

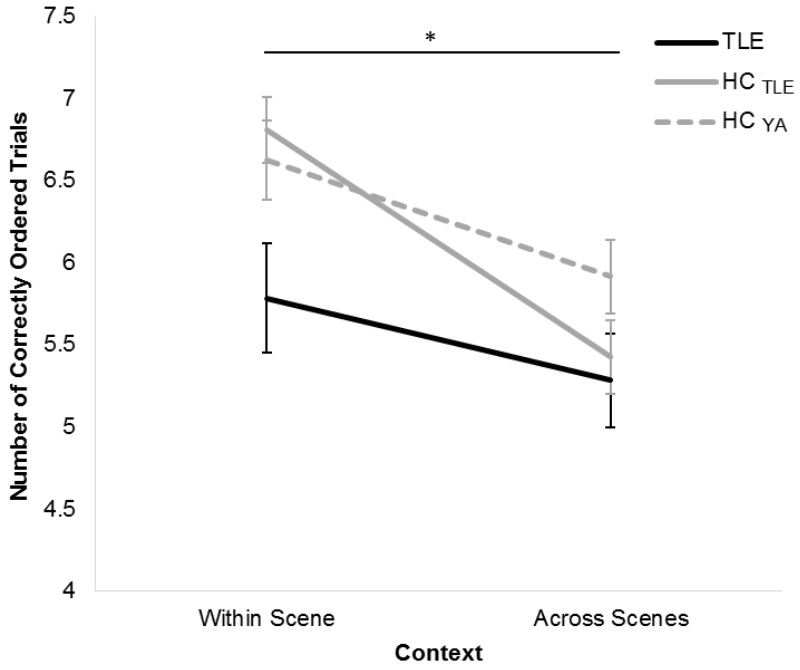
4.2.4 Influence of shared context on timeline performance

According to past literature, sequence memory is improved when items are presented within versus across event boundaries (DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011; Heusser et al., 2018; Zwaan, 1996), and items presented in the same context are remembered to have occurred closer together in time than items presented across event boundaries (Ezzyat & Davachi, 2014; Faber & Gennari, 2015). In the healthy-validation study, participants were significantly better at correctly ordering still frames sampled from the same scene compared to still frames sampled across one or more event boundaries (operationalized as scene changes; $Z = -2.17$, $p = .030$, $\eta^2 = .10$). Additionally, the interval between still frames sampled within a scene was rated as significantly shorter than the interval between still frames sampled across scenes ($Z = -4.29$, $p < .001$, $\eta^2 = .39$), despite the sampling interval for both conditions actually being identical. In other words, using the timeline-judgement task, we were able to replicate temporal-memory manipulations observed with other tasks.

Next, we sought to replicate the influence of shared context on temporal memory in epilepsy. Using a mixed model ANOVA, we investigated the effects of group (TLE versus HC_{TLE}) and context (within versus across scenes) on the number of correctly ordered trials. Just as in the healthy-validation sample, both groups demonstrated a context effect, such that participants were better able to correctly order still frames when they were sampled within a single scene than across scenes, $F(1,32) = 20.68$, $p < .001$, $\eta^2 = .36$. A trending, though non-significant, group effect suggests that the TLE participants had somewhat more difficulty in correctly ordering still frames overall in the trials used for this analysis (i.e., those items with still frames sampled across intervals of 15 and 30 s; $F(1,32) = 3.31$, $p = .078$, $\eta^2 = .09$). In addition, the effect of context differed significantly between groups, $F(1,32) = 5.15$, $p = .030$, $\eta^2 = .09$. Simple-effect analysis revealed that the TLE and control groups were similarly able to correctly order trials when still frames were sampled across scenes ($p = .749$); however, the control group outperformed the TLE group when still frames were sampled within the same scene

($F(1,32) = 7.41$, $p = .010$, $\eta^2 = .19$). Repeating this ANOVA with the observed interval (selected by the participant) as the dependent variable, we were again able to replicate the effect of context observed in the healthy-validation study; participants judged still frames sampled from the same scene as closer together than still frames sampled across scenes, $F(1,32) = 59.25$, $p < .001$, $\eta^2 = .65$. A significant group difference ($F(1,32) = 4.67$, $p = .033$, $\eta^2 = .13$) suggests the TLE group rated still frames as occurring somewhat further apart than the control group. The interaction of group and context was not significant, $p = .716$. The results are displayed graphically in Figure 8.

A. Effect of context and group on correctly ordered trials



B. Effect of context and group on interval estimation

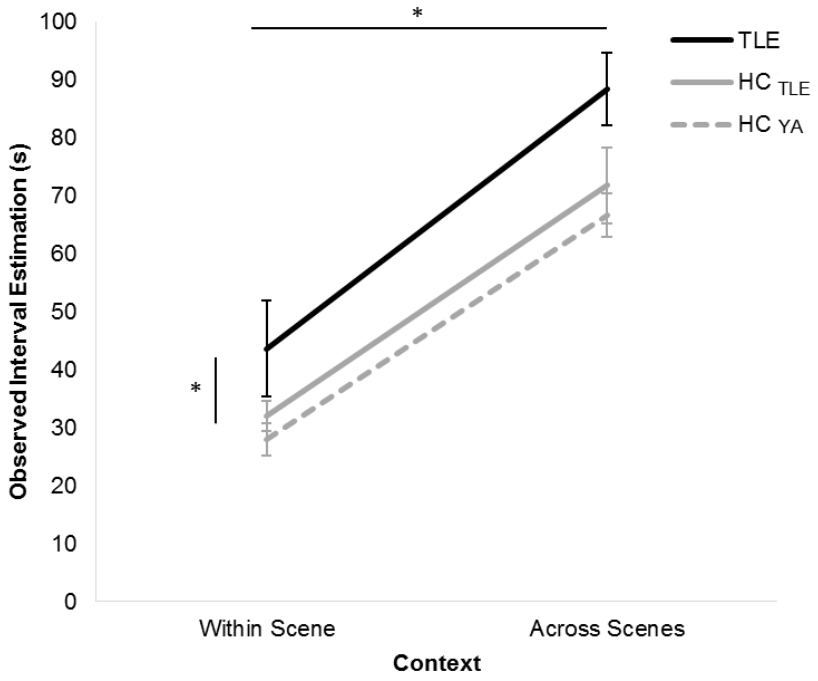


Figure 8. Effect of still-frame context and group on timeline performance. Error bars represent standard error of the mean. TLE = temporal lobe epilepsy sample; HC$_{TLE}$ = healthy control participants – matched to epilepsy sample; HC$_{YA}$ = healthy control participants – young adults.

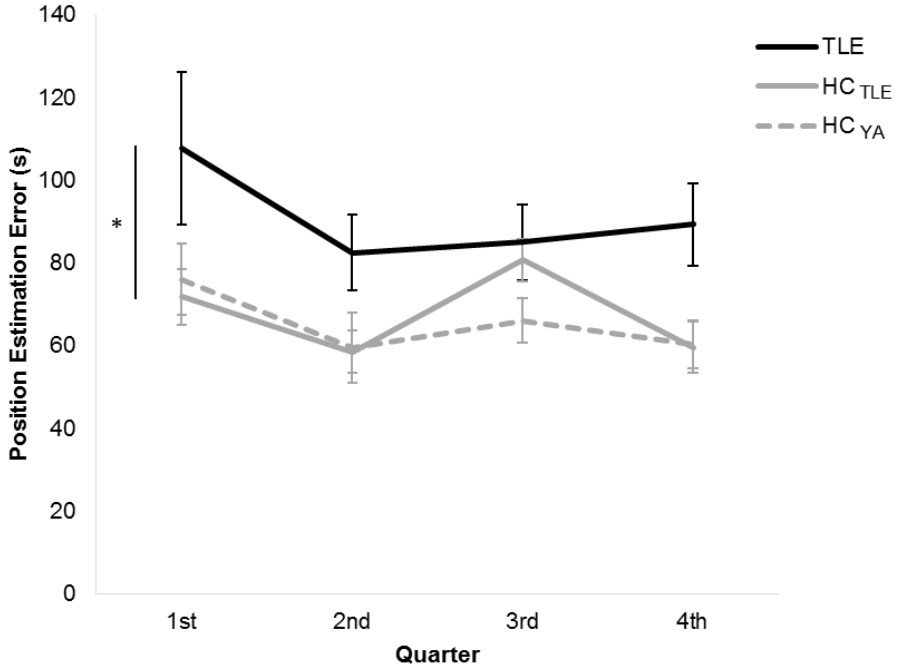
* significant at the .05 level

4.2.5 Influence of sampling parameters on timeline performance

In the HC$_{YA}$ sample, accuracy of position estimation was significantly different across the four quarters of the movie clip ($\chi^2(3) = 8.85$, $p = .031$, *Kendall's W* = .12). Follow-up pairwise comparisons revealed significantly greater position estimation error for the first versus second quarter ($Z = -2.06$, $p = .040$, $\eta^2 = .09$). Accuracy of interval estimation was significantly affected by the true interval at which the still frames were sampled, $F(2,48) = 35.91$, $p < .001$, $\eta^2 = .61$. Interval estimation error by true interval length followed a quadratic trend ($F(1,23) = 24.30$, $p < .001$), such that estimation accuracy became increasingly worse when the true sampling interval exceeded 60 s.

The influence of sampling parameters was also investigated in the TLE and HC_{TLE} samples. Using a mixed model ANOVA, we investigated the influence of the quarter of the movie (from which the still frame was sampled) and group on position estimation error. Position estimation error marginally differed across the four quarters of the movie clip, $F(2,76) = 2.50$, $p = .080$, $\eta^2 = .07$. A significant main effect of group revealed better position-estimation accuracy in the control group, $F(1,32) = 8.62$, $p = .006$, $\eta^2 = .21$. No interaction of quarter and group was observed, $p = 0.170$. A second mixed model ANOVA explored the influence of sampling interval and group on interval estimation error. Similar to the healthy validation, there was a main effect of interval ($F(3,84) = 45.53$, $p < .001$, $\eta^2 = .57$), with worse interval estimation most apparent at longer true intervals. A main effect of group demonstrated better interval estimation accuracy in the control group, $F(1,32) = 5.03$, $p = .032$, $\eta^2 = .14$. The interaction of interval and group was not significant, $p = .072$. These results are displayed graphically in Figure 9.

A. Effect of sampling quarter and group on position estimation



B. Effect of sampling interval and group on interval estimation

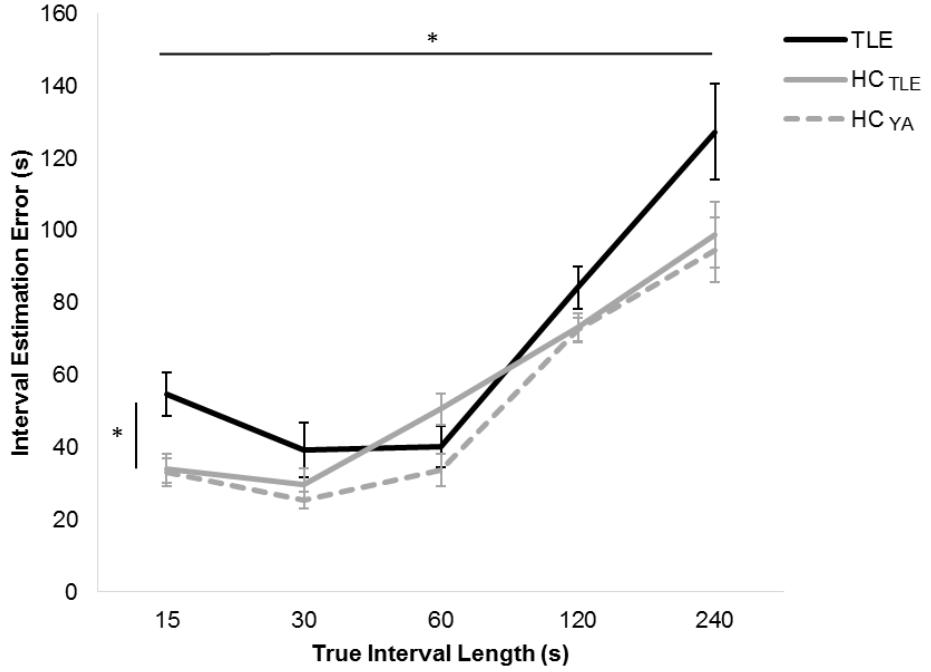


Figure 9. Effect of item-sampling parameters and group on timeline performance. Error bars represent standard error of the mean. TLE = temporal lobe epilepsy sample; HC_{TLE} = healthy control participants – matched to epilepsy sample; HC_{YA} = healthy control participants – young adults.

* significant at the .05 level

4.2.6 Logistic regression

Binary logistic regression models were used to predict group membership (TLE versus control) from standardized and temporal-memory measures. Among the standardized measures, the Doors test showed greatest sensitivity to group membership. For temporal-memory measures, composite measures were used instead of individual measures to reduce the number of regressors and maximize sensitivity to group. In the first set of hierarchical regression models, Doors was entered first as the standardized clinical measure, followed by the recall & timeline composite. The first model based on Doors alone significantly predicted group ($\chi^2(1) = 10.57$, $p = .001$), explained 33% of the variance between groups (Nagelkerke R^2), and correctly classified 71% of participants.

Addition of the recall & timeline composite significantly improved prediction ($\chi^2(1) = 7.04$, $p = .008$). This new model significantly predicted group ($\chi^2(2) = 17.61$, $p < .001$), explained 50% of the variance between groups, and correctly classified 79% of participants. In this model, the odds of falling in the control group were increased by 35.92 times (95% CI: [1.52, 848.16]) with every unit increase in the recall & timeline composite ($B = 3.58$, Wald $\chi^2 = 4.93$, $p = 0.026$). The inclusion of the recall & timeline composite rendered Doors an insignificant predictor ($p = .147$), suggesting a degree of shared variance, despite low multicollinearity among these two predictors ($VIF = 1.0$). Comparing TLE participants who were correctly classified ("true positives") to those who were not ("false negatives"), these subgroups show similar proportions of individuals with left versus right TLE ($\chi^2(1) = 1.66$, $p = .198$) and with probable MTS versus equivocal or negative findings on clinical MRI ($\chi^2(1) = .93$, $p = .334$), although sensitivity to differences is poor because sample sizes are small. However, among the 10 individuals who were correctly identified as TLE participants, it is notable that three showed no evidence of structural abnormality on clinical MRI and one showed equivocal evidence of MTS. Therefore, despite lacking clear evidence of an epileptogenic lesion in the temporal lobe, their performance on the Doors and temporal-memory measures could be used to distinguish them from the HC_{TLE} sample.

To investigate whether either temporal-memory composite alone could be modeled with Doors to produce similar predictive value to the recall & timeline composite, we repeated the above hierarchical regression using the recall composite and timeline composite separately. The model based on Doors and the timeline composite significantly predicted group ($\chi^2(2) = 16.52$, $p < .001$), with the addition of the timeline composite significantly improving prediction ($p = .015$). This model explained 48% of the variance in group, and correctly classified 79% of participants. In contrast, addition of the recall composite only marginally improved prediction of group above and beyond the model with Doors alone ($p = .070$). This model showed a classification accuracy of 76%.

Finally, a model based on the timeline composite alone was tested. This model significantly predicted group ($\chi^2(1) = 12.12$, $p = .001$), explained 37% of the variance between groups, and correctly classified 82% of participants.

Classification statistics for the four models tested are displayed in Table 13. Despite similar success in overall classification across models, the sensitivity (probability of labeling a TLE participant as TLE), specificity (probability of labeling a HC_{TLE} participant as HC_{TLE}), positive predictive value (probability of those labeled as TLE actually being TLE), and negative predictive value (probability of those labeled as HC_{TLE} actually being HC_{TLE}) of these models differ to some degree. Overall, the model based on timeline alone yielded the strongest classification statistics.

Table 13. Classification statistics based on logistic regression models to predict group (TLE vs. HC_{TLE}).

Model	Classification Accuracy (%)	Sensitivity	Specificity	PPV	NPV
1. Doors, 2. Recall & Timeline Composite	79	.67	.87	.77	.80
1. Doors, 2. Recall Composite	76	.60	.87	.75	.77
1. Doors, 2. Timeline Composite	79	.73	.83	.73	.83
1. Timeline Composite	82	.75	.86	.80	.83

Note. The models represent hierarchical logistic regression models. Classification statistics are based on a cut-off value of .5 applied to the predicted group probabilities generated by the regression equations. TLE = temporal lobe epilepsy sample; HC_{TLE} = healthy control participants – matched to epilepsy sample; PPV = positive predictive value; NPV = negative predictive value.

4.3 Discussion

The goal of this study was to investigate the potential utility of a novel, naturalistic, lab-based test of temporal memory in the evaluation of individuals with TLE. As a first step, we compared the TLE group and demographically matched controls on temporal-memory performance, and the TLE group showed significantly worse performance on these novel measures. In our brief battery of standardized memory tests, the only measure to significantly distinguish groups was the Doors test, with additional measures (CALT, RVDLT) showing marginal effects of group. Group accounted for a similar proportion of variance (~30%) in Doors performance as in performance on temporal-memory measures. This promising finding suggests that the temporal-memory measures are sensitive to cognitive difficulties manifest in TLE

We also investigated the psychometric properties of these measures in our healthy and TLE samples. Since the recall task used a subjective scoring system, we investigated its interrater reliability. The resultant correlation coefficients were “fair” to “excellent” by clinical standards (Cicchetti, 1994), but still somewhat lower than we anticipated based on similar investigations (St-Laurent et al., 2011). Another measure of reliability, a form of internal consistency, could be gleaned from the correlations performed within each of the temporal-memory tasks. All pairwise correlations among recall measures and most among timeline measures demonstrated moderate to strong relationships. High intra-test correlations provide support for the construction of composite measures for these tests. The significant weak to moderate correlations between recall and timeline measures suggest some imperfect degree of overlap in the construct(s) that they assess.

The unintuitive observation that recall sequencing errors was directly proportional to other measures of optimal performance on the two tests seems to suggest that producing more details puts the participant at risk of producing more details out of sequence. This tendency was exaggerated in the TLE group, as they demonstrated significantly stronger correlations among sequencing errors and other recall measures compared to controls. In our group comparisons, just as in the paper by St-Laurent and colleagues (2011) on

which our scoring system was based, we found that both groups had a low base rate of sequencing errors and that these rates were not significantly different between the TLE and matched control group. As well, sequencing errors did not correlate with any standardized cognitive measures. In summary, total sequencing errors appeared to trend with the number of details recalled but was not itself a sensitive measure of episodic memory or temporal-lobe dysfunction.

To investigate convergent and discriminant validity, temporal-memory measures were also correlated with standardized cognitive tests. Among recall measures, correlations with standardized memory measures did not survive FDR correction. On the other hand, timeline measures demonstrated moderate to strong correlations with tests of verbal memory (Logical Memory, CVLT), visual memory (RCFT, RVDLT, Doors), and verbal intelligence (Vocabulary). No correlations were observed with a measure of verbal recognition memory (Names), verbal recall memory (RAVLT), self-report episodic memory (SAM Episodic) or nonverbal problem solving (Matrix Reasoning). Based on these results, we can be reasonably confident that the timeline measures are tapping into an episodic memory construct that includes both visual and verbal encoding/recall modalities.

Construct validity was also investigated by replicating the effects of shared context/event boundaries on temporal memory as demonstrated by other paradigms. Specifically, the timeline-judgement task lends itself to studying these effects as the two still frames presented in each item were sampled either within a scene (shared context) or across scene changes (event boundaries). From previous studies, we would expect ordering of still frames to be easier when they are sampled within versus across scenes (DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011; Heusser et al., 2018; Zwaan, 1996), and that interval estimation between still frames sampled from the same scene would be shorter than still frames sampled across scene changes (Ezzyat & Davachi, 2014; Faber & Gennari, 2015). The control groups performed as expected, suggesting that we were capturing effects similar to those observed in other, less naturalistic, tasks. The TLE

group demonstrated an effect of shared context on interval estimation (despite rating intervals as significantly longer than controls), but the benefit of shared context to sequence memory was less apparent in this group. Based on research conducted in healthy participants demonstrating a link between hippocampal activation and temporal separation across event boundaries (Baldassano et al., 2017; Ben-Yakov et al., 2013; Ben-Yakov & Henson, 2018; Deuker et al., 2016; Dubrow & Davachi, 2014; Ezzyat & Davachi, 2011, 2014), it follows that reduced hippocampal integrity would dull the effects of event boundaries on the temporal organization of memory, causing more consistent performance when items are sampled in shared and non-shared contexts. Indeed, Zacks and colleagues (2006) showed that another clinical sample with compromised hippocampal integrity, individuals with mild cognitive impairment, made aberrant judgements of event boundaries that coincided with impaired sequence memory.

In addition to whether the still frames were sampled within or across scenes, the effects of other sampling parameters were investigated to improve our understanding of factors that affect performance. With respect to position estimation, the quarter from which each still frame was sampled affected performance, with lowest accuracy observed for items sampled from the first quarter of the movie. With respect to interval estimation, as the true interval at which two still frames were sampled increased, the participants' estimates of that interval grew less accurate. Despite generally lower accuracy in the TLE group, they showed reasonably similar effects of these sampling parameters on performance. Of course, narrative elements of the movie (e.g., the number and salience of events within a particular quarter or interval) may also be expected to influence memory for those time periods. Knowledge of the factors that improve or hinder performance is useful in designing new tests for clinical application, as they can be used to improve sensitivity and to manipulate the difficulty of the test to avoid ceiling and floor performance effects.

Using logistic regression, we investigated the value of cognitive variables in predicting group membership (TLE versus control). The composite measure composed of recall and timeline measures improved upon the prediction and classification accuracy

demonstrated by the Doors test alone. We cannot, based on this one finding, argue for the inclusion of temporal-memory measures in neuropsychological batteries to enhance detection of temporal-lobe dysfunction in refractory epilepsy. However, the added predictive value of temporal measures above another strong predictor provides preliminary support that temporal measures may capture temporal-lobe dysfunction in ways that complement current standardized testing.

Performance on temporal-memory measures did not distinguish TLE participants with right-hemisphere foci from those with left-hemisphere foci. In addition, TLE subgroups based on laterality and evidence of structural lesion were equally likely to be misclassified by the regression equation based on Doors and temporal-memory performance. Therefore, the localizing information that can be discerned from task performance is limited. Failure to find a laterality effect is not surprising considering how these effects are typically captured in the clinical neuropsychological assessment: tests of verbal memory are thought to be more sensitive to left temporal dysfunction and tests of visual memory to right (Gleissner et al., 1998; Glosser et al., 1995; Helmstaedter et al., 1991; Hermann et al., 1997; Kim et al., 2003; Milner, 1972; Sass et al., 1995). In our paradigm, we tested memory for a dually-encodable audiovisual stimulus using verbal (recall task) and visual (timeline task) recall methods, and performance was correlated with standardized tests of verbal and visual memory. We may be underpowered in testing a laterality effect (comparing 6 left to 9 right TLE); however, St-Laurent and colleagues (2011), whose scoring system for autobiographical memories we adapted for the movie in the present study, also detected no group difference on temporal resolution or sequencing measures in their left (n = 14) versus right (n = 11) TLE participants.

We refer to the novel tasks as measures of temporal memory. However, we must consider whether temporal memory is what we are capturing. First, since we used a naturalistic stimulus, temporal context may not be directly encoded but rather inferred from other remembered details according to established (learned) schemas. Stories like the one depicted in the clip adhere to narrative rules that guide its progression, such that a

participant would be unlikely to begin retelling a story at its climax, for example. Story memory can also be bolstered by semantic knowledge, such that if the participants remember a temporally higher-order detail like “the boy was unpacking for his uncle,” they are more likely to recall temporally clustered details like “he opened the suitcase, he took out some clothes, and placed them in a drawer” based on their knowledge of what unpacking typically entails. Second, based on how our composite measures were computed to maximize group differences, the recall composite simply represented the total number of details recalled at all levels of temporal resolution. An alternative composite measure of memory for temporal context may have relied on some ratio of details at higher and lower temporal resolution. Indeed, St-Laurent and colleagues (2011) found that participants with TLE produced fewer details at higher temporal resolution (clustered details) but not at lower temporal resolution (higher order details). However, in our sample, participants with TLE produced fewer details at both higher and lower temporal resolution, likely reflecting the restricted time window of the eight-minute movie compared to an autobiographical memory of several hours. Furthermore, St-Laurent and colleagues (2011) note that producing fewer details at higher temporal resolution may not reflect a specific temporal-memory deficit, but rather a general lack of “vividness” that has also been demonstrated in other modalities (e.g., perceptual details; St-Laurent, Moscovitch, et al., 2014). Similarly, Brunec and colleagues (2017) found that although sequencing information could be gleaned from contextually impoverished memories (rated as “familiar”), duration estimation relied on more vividly reexperienced episodes (rated as “recollected”). Thus, the measures we refer to as temporal-memory measures may also reflect the precision or vividness with which episodic memories are recalled.

As possible clinical tools, the recall and timeline measures have strengths and weaknesses. We have demonstrated preliminary evidence of their sensitivity to temporal-lobe dysfunction, but performance on these tests does not seem to offer information about seizure lateralization or correlate with evidence of medial temporal sclerosis, at least in the small sample tested here. However, our results also suggest that the tests may be

sensitive to temporal-lobe dysfunction even in so-called “MRI negative” cases, with no or limited evidence of a lesion on structural MRI, for whom a more intensive presurgical investigation is often needed. Like many neuropsychological tests, these measures would also be expected to show linguistic and cultural bias. We did not test these effects explicitly, however, all participants were native English speakers and most spoke English as their first language. Perhaps not surprisingly, a measure of verbal intelligence (Vocabulary) was correlated with task performance, suggesting that difficulties with verbal comprehension or expression may hinder performance for reasons unrelated to a memory deficit.

As individual tasks, the recall and timeline measures also showed relative strengths and weaknesses with respect to clinical application. From a practical perspective, to learn and implement the recall administration and scoring guidelines would be cumbersome, whereas the timeline task was administered electronically, and outcome measures were computed automatically. Another practical advantage of the timeline task is that the correct ordering measure can be used as a coarse indicator of test-taking effort, since above-chance responding can be a prerequisite for analysis of timeline data. Considering the psychometric properties reported above, construct validity could be reasonably well established for the timeline measures but not the recall measures. Even if temporal memory is not the intended target construct for any standardized tests, we would still expect a certain degree of correlation with other episodic memory tasks, which was not observed for recall measures. Furthermore, unlike the timeline task, recall only marginally improved the prediction of group on top of the predictive value of the Doors test, i.e., its contribution to group prediction was largely redundant when the contribution of the Doors test had already been taken into account. In fact, group prediction based on the timeline task alone showed the strongest classification statistics of all of the models tested. In summary, of these two measures, the timeline task has more favourable practical and psychometric qualities. A caveat to interpreting these tasks in isolation is that they were always administered in the same order (recall before timeline). Thus, to

confirm the value of the timeline task, the above analyses would ideally be replicated in an independent sample who completed only the timeline task.

We have demonstrated the potential utility of measures that assess memory for the temporal features of naturalistic, audiovisual stimuli. These temporal-memory measures appear to demonstrate sensitivity to temporal-lobe dysfunction in TLE and may capture aspects of functioning that are not currently measured as part of standardized memory testing. The present findings are promising, and suggest that the inclusion of tasks like these in the neuropsychological assessment of individuals with TLE may enhance detection of temporal-lobe dysfunction and provide novel information for surgical planning.

Chapter 5

5 Movie-driven fMRI and subsequent memory testing in temporal lobe epilepsy

When temporal lobe epilepsy (TLE) is refractory to antiepileptic drugs, a surgical resection of the epileptogenic tissue offers an effective alternative to control seizures (Engel, 2001). To balance the benefit of seizure reduction with the potential cost of disrupting essential cognitive and sensorimotor functions, surgical candidates must undergo extensive investigations to plan an optimal surgical approach. These investigations typically include structural neuroimaging, video-electroencephalography (EEG) monitoring, and neuropsychological assessment; however, when standard assessment techniques are inconclusive, a variety of additional tools may be called upon to provide extra lateralizing or localizing information (Datta & Loddenkemper, 2011). Convergence of findings across these multidisciplinary investigations enhances certainty that the epileptic focus has been successfully identified, potentially improving postsurgical outcomes.

Rather that inferring relationships across assessment tools, a protocol that allows direct comparison of different sources of information may provide more convincing evidence of brain-behaviour relationships. In the current study, we conducted a functional magnetic resonance imaging (fMRI) scan while participants freely viewed a movie stimulus, and then tested their memories for the movie. Our goal was to quantify the relationship between neuroimaging measures derived from movie-driven fMRI and cognitive measures derived from the movie-memory test. Both structural (volume) and functional (connectivity) measures of medial temporal-lobe integrity were obtained.

Structural neuroimaging, primarily magnetic resonance imaging (MRI), is a standard investigation in presurgical epilepsy used to identify anatomical abnormalities that reflect an underlying etiology (Rosenow & Lüders, 2001). For example, mesial temporal sclerosis (MTS), consisting of cell loss and gliosis in the hippocampal formation, can be

identified on MRI in 50 to 70% of individuals with refractory TLE (Briellmann et al., 2002; Falconer, Serafetinides, & Corsellis, 1964) and the evidence of a structural lesion like MTS on presurgical MRI is predictive of significantly better postsurgical seizure relief compared to non-lesional cases of epilepsy (Tellez-Zenteno et al., 2010). Studies using MRI volumetry consistently show relatively reduced left hippocampal volume in left TLE and right hippocampal volume in right TLE (Barnett et al., 2015; Berkovic et al., 1991; Bernasconi et al., 2003; Doucet et al., 2016; Fuerst et al., 2001; Lencz et al., 1992; Mechanic-Hamilton et al., 2009).

Since hippocampal volume is thought to reflect structural integrity, a number of investigators have studied the relationship between reduced hippocampal volume and memory impairment in TLE. Left hippocampal volume in left TLE has been consistently linked to verbal-memory abilities, whereas the link between right hippocampal volume and non-verbal memory in right TLE is not so apparent (Alessio et al., 2006; Doucet et al., 2016; Glikmann-Johnston et al., 2008; Lencz et al., 1992; Rausch & Babb, 1993). Other studies use the combination or asymmetry of hippocampal volumes in TLE to investigate these structure-function relationships. For example, in their mixed group of right and left TLE participants, Barnett and colleagues (2015) demonstrated a relationship between left and right hippocampal volume asymmetry and verbal and nonverbal memory asymmetry. Reminger and colleagues (2004) did not find a significant association between hippocampal volume asymmetry and either verbal- or visual-memory measures, but they did find an association between combined hippocampal volume and a standardized measure of delayed story recall. This finding is consistent with that of Stoub and colleagues (2019), who showed that left and right hippocampal volume individually correlated with the same measure of story recall. Thus, both individually and in combination, hippocampal integrity (reflected by MRI-derived volumes) has been linked to memory ability (reflected by standardized memory measures, particularly verbal memory).

Functional neuroimaging, especially fMRI, is also used in many epilepsy surgery centres to inform surgical planning. Task-based fMRI is used to investigate the neural correlates of language and memory functions to (a) lateralize hemispheric dominance for language, (b) assist with seizure localization, and (c) contribute to the prediction of postoperative cognitive changes (Benjamin et al., 2018; Limotai & Mirsattari, 2012; McAndrews, 2014; McAndrews & Cohn, 2012). The importance of mapping language and memory functions in TLE is underscored by a higher likelihood of atypical (right or bilateral) hemispheric dominance for language compared to the general population (Branch et al., 1964; Duchowny et al., 1996; Hamberger & Cole, 2011; Möddel et al., 2009; Rausch & Walsh, 1984) with potential concomitant reorganization of memory (Alessio et al., 2013; Gleissner et al., 2002; Powell et al., 2007; Richardson et al., 2003; Seidenberg et al., 1997). Although many language-fMRI paradigms are considered to have adequate validity for clinical use to lateralize language dominance (e.g., high concordance with direct cortical stimulation and Wada testing; Janecek et al., 2013; Szaflarski et al., 2008), memory-fMRI paradigms yield mixed results and have not become standard of care (McAndrews, 2014). Nonetheless, several investigators have demonstrated a link between hippocampal activation during task-based fMRI and memory abilities pre- and postsurgically in TLE (Binder et al., 2008; Powell et al., 2008; Rabin et al., 2004). However, since task-based fMRI paradigms often involve no or limited overt behavioural responses, the relationship between task activation and memory abilities is often inferred across different investigations (i.e., interpreting fMRI activation in light of memory performance on standardized measures; Baxendale & Thompson, 2010).

Resting-state fMRI (rs-fMRI) has been investigated as a complement to task-based fMRI for mapping functional brain networks. Functional connectivity analyses based on rs-fMRI have an intuitive application in epilepsy; despite the emphasis placed on localization of a seizure focus, epilepsy itself can be conceptualized as a disorder of networks (Engel et al., 2013; Spencer, 2003). In other words, focal abnormalities can have downstream consequences in functionally connected regions. Rs-fMRI studies in TLE have focused on abnormal hippocampal connectivity, and consistently document

altered (both increased and decreased) connectivity across nodes of the default-mode network (DMN), which include medial temporal, lateral parietal, posterior cingulate, and medial prefrontal regions (Cataldi, Avoli, & De Villers-Sidani, 2013; de Campos, Coan, Lin Yasuda, Casseb, & Cendes, 2016; Doucet, Osipowicz, Sharan, Sperling, & Tracy, 2013; Holmes et al., 2014; James, Tripathi, Ojemann, Gross, & Drane, 2013; Liao et al., 2011; McCormick et al., 2014; McCormick, Quraan, Cohn, Valiante, & McAndrews, 2013; Pittau, Grova, Moeller, Dubeau, & Gotman, 2012; Voets et al., 2014, 2012). As well, inter- and intra-hemispheric effects can be observed in unilateral TLE, including weakening of connections in the affected hemisphere and strengthening in the contralateral hemisphere (Bettus et al., 2009; Maccotta et al., 2013; Su, An, Ma, Qiu, & Hu, 2015), and altered (increased and decreased) coupling of bilateral hippocampi and other homologous regions in the temporal lobes (Maccotta et al., 2013; Morgan, Rogers, Sonmezturk, Gore, & Abou-Khalil, 2011; Pittau et al., 2012; Tracy et al., 2014).

Only a handful of these investigations relate hippocampal connectivity to memory abilities to assess whether specific functional alterations might have cognitive consequences. McCormick and colleagues (2014, 2013) demonstrated a relationship between alterations of DMN connectivity and material-specific memory in left and right TLE. Other studies reveal alterations in interhemispheric or contra-lesional connectivity that may reflect recruitment of the non-lesional hemisphere. In left TLE, Holmes and colleagues (2014) showed that increased connectivity between the left hippocampus and right precuneus and inferior parietal areas was associated with better verbal memory, whereas increased connectivity between the left hippocampus and left precuneus and inferior parietal areas was associated with worse verbal memory. In right TLE, Doucet and colleagues (2013) showed that reduced connectivity between left medial temporal structures and medial frontal cortex was associated with reduced delayed nonverbal recall.

Beyond rs-fMRI, other continuous acquisition paradigms under “active” or naturalistic conditions, like watching a movie clip, reveal different patterns of functional connectivity

(Bartels & Zeki, 2005; Betti et al., 2013; Geerligs, Rubinov, Cam-CAN, & Henson, 2015; Vanderwal et al., 2017), and may offer several advantages over rs-fMRI. First, rs-fMRI can be used to identify functionally coherent networks but not the actual functions subserved by these networks. In contrast, networks that are stimulus driven can be directly related to stimulus features to elucidate underlying cognitive processes. Reversing this logic, stimulus features can also be selected to modulate activity in networks of interest. For example, to investigate hippocampal connectivity in the present study, we used a movie stimulus with features known to modulate hippocampal activity, such as novelty to the viewer (Kumaran & Maguire, 2009; Tulving et al., 1996), perceptual richness (Robin & Moscovitch, 2017), and numerous event boundaries (Baldassano et al., 2017; Ben-Yakov et al., 2013; Ben-Yakov & Henson, 2018). In addition, movie viewing compared to rest may produce more sensitive functional-connectivity analyses by reducing head motion (Centeno et al., 2016; Huijbers et al., 2017), and increasing reliability of activation patterns within (Wang & Diana, 2016) and across participants (Hasson et al., 2004).

Naturalistic stimulation paradigms can also be combined with memory testing for the stimulus to directly investigate the relationship between observed activation and memory ability. For example, Hasson and colleagues (2008) asked participants to watch a 27-minute television episode while undergoing fMRI, and after three weeks, participants completed a memory test for the movie, designed such that each question could only be answered by recalling a 20-second segment of the episode. They then identified regions that showed stronger intersubject correlations (reflecting the consistency of activation over time across viewers) for remembered versus forgotten segments of the episode, which included the parahippocampal gyrus, superior temporal gyrus, anterior temporal poles, and the tempero-parietal junction. Lositsky and colleagues (2016) showed that retrospective judgements of the length of an interval between two radio clips was associated with the extent of fMRI pattern change (in medial temporal and prefrontal regions) between the two clips at encoding. Ben-Yakov and colleagues (2011; 2013, 2014) related memory for the gist of movie clips (4 to 16 s in length) to the extent of

hippocampal activation at the offset of clips. These studies demonstrate how the rich content of naturalistic stimuli can be leveraged to investigate the cognitive and neural processes underlying everyday memory abilities.

Other investigations (e.g., Furman et al., 2012), including several in TLE (Bonnici et al., 2013; St-Laurent et al., 2016), have scanned participants while they retrieved, rather than encoded, a movie stimulus. St-Laurent and colleagues (2016) showed that retrieval of perceptual details (greater for movies and autobiographical memories than for narrative scripts) was associated with increased activation in a number of regions including the right hippocampus, and that individuals with right TLE showed an attenuated perceptual-richness signal in these regions, consistent with their difficulty in retrieval perceptual details (St-Laurent, Moscovitch, et al., 2014). Bonnici and colleagues (2013) asked participants to recall movie clips of everyday events (7 s in length) while undergoing fMRI. They trained a classifier to learn the patterns of brain activity associated with each memory, and found predictable patterns of activity in the contra-lesional but not the sclerotic hippocampus of individuals with refractory TLE (potentially demonstrating functional reserve).

In the current study, we asked participants to watch an engaging eight-minute audiovisual film clip while undergoing an fMRI scan. Outside of the scanner, participants completed a memory test for the movie. We then tested the associations between structural (hippocampal volume) and functional (hippocampal connectivity) measures derived from neuroimaging and cognitive measures derived from the movie-memory test as well as standardized cognitive measures. Our goal was to investigate whether this novel movie-based fMRI/memory assessment paradigm could provide meaningful information about the hippocampal network and associated memory abilities in TLE.

5.1 Methods

5.1.1 Participants

Three samples of participants were included in the present study and are summarized in Table 14. The TLE group was comprised of 19 participants recruited from the Adult Epilepsy Service at London Health Sciences Centre (London, Ontario, Canada), who were undergoing presurgical evaluation for a temporal-lobe resection to control seizures. For all participants, one or more potential temporal-lobe foci had been identified on electroencephalography (EEG), magnetic resonance imaging (MRI), and/or intracranial EEG (iEEG). None had undergone previous brain surgery.

Table 14. Group characteristics.

	TLE	HC_{TLE}	HC_{YA}
N	19	24	24
Sex (F:M)	9:10	14:10	14:10
Age (*M* ± *SD*)	33.79 ± 12.25	35.67 ± 14.89	23.17 ± 3.24
Years of Education (*M* ± *SD*)	12.79 ± 2.25	13.74 ± 1.76	15.63 ± 2.32
Handedness (R:L)	18:1	19:5	24:0
Seizure Lateralization	9R: 8L: 2BL	.	.
MRI evidence of MTS	5R: 5L: 2BL: 7 none		
Years since onset (*M* ± *SD*)	17.47 ± 16.33	.	.

Note. Seizure lateralization was based on EEG or iEEG evidence of seizures originating from one or both temporal lobes. TLE = temporal lobe epilepsy sample; HC_{TLE} = healthy control participants – matched to epilepsy sample; HC_{YA} = healthy control participants – young adults; F = female; M = male; R = right; L = left; BL = bilateral; MRI = magnetic resonance imaging; MTS = mesial temporal sclerosis.

Two healthy control (HC) samples were recruited. A group of 24 control participants demographically matched to the TLE sample (HC_{TLE}) were recruited from the wider London community. Specifically, the HC_{TLE} group was matched to the TLE group on age (p = .826), years of education (p = .070), and sex distribution (p = .474). A sample of healthy young adults (HC_{YA}) was recruited to investigate the psychometric properties of the movie memory test (the results of which are discussed in Chapters 3 and 4), and to

increase the overall sample size for hypothesis testing. The HC_{YA} sample consisted of 24 participants recruited predominantly from Western University's undergraduate population. All healthy control participants reported no history of neurological or psychiatric illness. They were also native English speakers (i.e., learned English before age 5), with no hearing difficulties (based on self-report) and normal or corrected-to-normal vision.

5.1.2 Procedure

Participants underwent a functional magnetic resonance imaging (fMRI) scan and subsequently completed a memory test for the movie. For various reasons (technical issues, time limitations), three of 19 TLE participants and one of 24 HC_{TLE} participant completed the scan but did not complete the movie memory test. Therefore, analyses based on fMRI-derived measures alone include the full sample, whereas analyses that integrate fMRI and cognitive testing exclude the four participants without movie memory test data. The fMRI scan was acquired according to the standards of EpLink, the epilepsy research program of the Ontario Brain Institute. The acquisitions of interest for the present study were a T1-weighted structural scan and T2*-weighted functional scan during which the movie was played. The movie-driven functional scan was always the final acquisition in the scanning protocol to minimize the time between viewing the movie clip and beginning the memory test (approximately 10 minutes). Prior to the fMRI scan, participants were notified that they would be asked to watch a short movie near the end of the hour-long scan, and they were instructed to pay attention and follow the plot of the movie as they would if they were watching any other movie or television show. They were not forewarned about the movie memory test. Participants were also administered a short battery of standardized neuropsychological tests, variably performed before the fMRI scan or after the movie memory test (for scheduling purposes).

5.1.3 Movie stimulus

Participants were shown an eight-minute long, black-and-white movie clip. The clip was edited from a 20-minute 1961 television episode entitled "Alfred Hitchcock Presents:

Bang! You're Dead" and preserved the original plot. This episode has been shown to elicit a spatially distributed pattern of reliable activation across participants in widespread areas of the cerebral cortex, including prefrontal areas that are not activated reliably with other clips (Hasson et al., 2010). Since it was originally broadcast in 1961, it also has the advantage of being novel to the participants to promote engagement with and interest in the clip.

5.1.4 MRI acquisition

Whole-brain imaging was performed on a 3T Siemens Magnetom Prisma scanner (Siemens Healthcare, Erlangen, Germany) with a standard 32-channel head coil. A T1-weighted anatomical image (TR = 2300 ms, TE = 2.98 ms, flip angle = 9˚, voxel size = 1 mm isotropic, FOV = 256 mm^2) using a 3D magnetization-prepared rapid acquisition with gradient echo (MPRAGE) pulse sequence was collected at the start of scanning. During movie viewing, a T2*-weighted functional scan (TR = 2000 ms, TE = 30 ms, flip angle = 75˚, voxel size = 3 mm isotropic, FOV = 192 mm^2) was collected using a gradient-echo echo-planar imaging sequence. The scanned volume included 33 slices of 3 mm thickness with an interslice gap of 25% collected in interleaved descending order.

The movie stimulus was presented using the Psychophysics toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997), projected on a screen behind the MRI bore, and reflected via a mirror mounted on the head coil. Participants were provided with insert earphones for sound delivery and, when necessary, MRI-compatible lenses to correct vision. Foam padding was used to restrict head motion.

5.1.5 Neuroimaging analysis

5.1.5.1 Preprocessing

Structural MRI data were reconstructed and anatomically segmented with FreeSurfer v5.3.0 (http://surfer.nmr.mgh.harvard.edu/), which has been shown to be sensitive to medial temporal lobe atrophy in TLE (Pardoe et al., 2009). Skull stripping was performed with FSL BET (Smith, 2002). From the automated subcortical segmentation results,

volumetric data were retrieved for all segmented regions in mm^3, and these values were divided by the total intracranial volume. Binary masks of the hippocampi (right and left) based on the FreeSurfer segmentation were created in MATLAB 2015a (Mathworks, Natick, MA, USA). In addition to whole left and right hippocampal masks, the individual masks were sectioned along the transverse plane into equal thirds; the anterior-most third is referred to as the anterior hippocampus and the posterior two-thirds as the posterior hippocampus, roughly corresponding to the hippocampal head and body/tail regions, respectively.

Functional MRI data were preprocessed with FSL FEAT version 3.14 (Woolrich, Ripley, Brady, & Smith, 2001), including removal of five initial dummy volumes, realignment to the middle volume (MCFLIRT; Jenkinson, Bannister, Brady, & Smith, 2002), brain extraction (BET), high-pass temporal filtering at .001 Hz, coregistration of functional to structural scans, and independent components analysis (ICA; MELODIC; Beckmann & Smith, 2004). FSL FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) was used to filter the results of the ICA for noise components. The FIX algorithm was trained based on manual classification (Griffanti et al., 2017) of a subset of 17 participants (all from the clinical-validation study; nine clinical participants, eight control participants), using the majority rating of three independent raters per subject. Structural and functional scans were spatially normalized in SPM12 (Wellcome Trust Centre for Neuroimaging, London UK) to the MNI 152 template.

Realignment parameters (derived from MELODIC) and tsdiffana plots (http://sourceforge.net/projects/spmtools) were visually inspected to ensure the quality of the fMRI data. No participant demonstrated unacceptable levels of variability over time (i.e., the y-axis of intensity/motion plots was comparable across participants), and signal intensity perturbations appear to affect less than 1% of the session's 241 volumes. Therefore, all participants and volumes were retained for analysis.

5.1.5.2 Functional connectivity

Functional connectivity analysis was performed using the CONN toolbox v18a (http://www.nitrc.org/projects/conn; Whitfield-Gabrieli & Nieto-Castanon, 2012). Prior to analysis, further denoising of the functional data was performed: global signal from white matter and cerebrospinal fluid masks, as well as individual realignment parameters, were regressed out. Functional coupling across analogous hippocampal regions in the two hemispheres was conducted using the six hippocampal masks (left and right whole, anterior, and posterior hippocampal masks) as regions of interest (ROIs). Anterior and posterior hippocampal ROIs were investigated separately based on previous evidence of functional dissociation and distinct connectivity (Robin & Moscovitch, 2017; Voets et al., 2014). In addition, hippocampal connectivity with the rest of the brain was investigated using the six hippocampal masks as source ROIs. For this ROI-to-ROI analysis, we assessed hippocampal connectivity with the CONN default ROIs based on atlas (Harvard-Oxford and AAL atlases) and resting-state network regions, with the whole hippocampal masks based on the FreeSurfer segmentation substituted for atlas-based hippocampal ROIs. Average timeseries data from each of the ROIs (six subject-specific hippocampal, 131 atlas, 32 network ROIs) were extracted for each individual participant. All possible bivariate correlations were performed, and the correlation coefficients were Fisher-transformed. In consideration of the numerous correlations tested, a false-discovery rate (FDR) correction was applied (Benjamini & Hochberg, 1995). Similarly, when group comparisons were performed, FDR correction was applied to control for multiple comparisons.

5.1.6 Movie-memory test

The memory test was composed of four sections, administered in the following order: recall, familiarity judgements, timeline judgements, and comprehension questions. Recall was orally administered, and the other three sections were administered on a laptop via Psychophysical toolbox. Sample items and scoring procedures are shown in Figure 10.

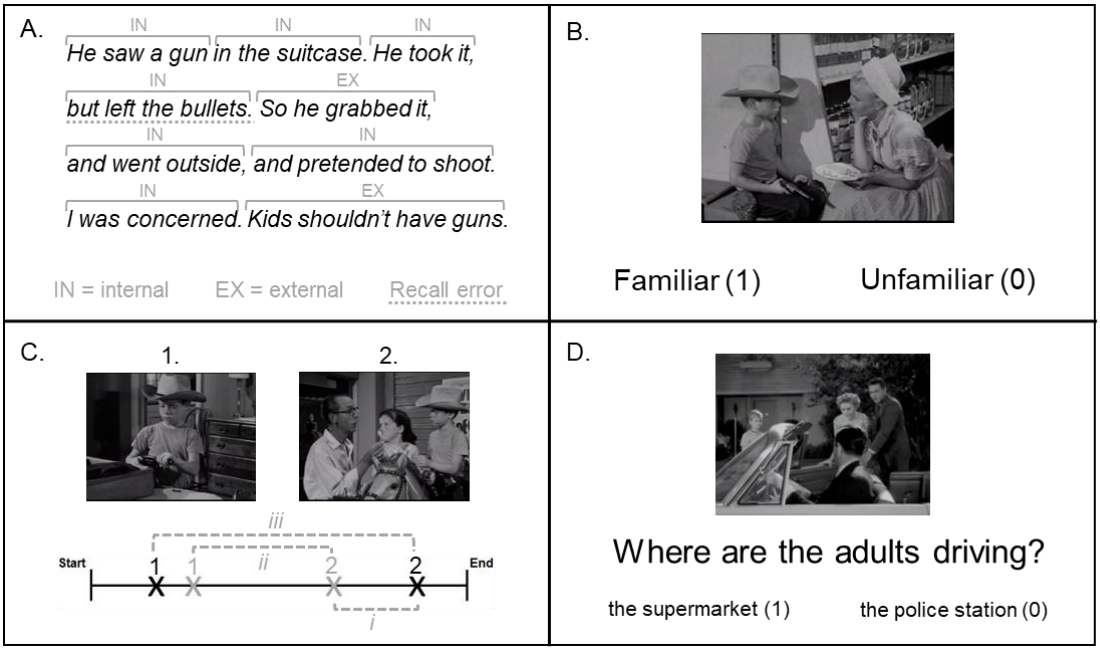


Figure 10. Sample items/segments from the movie-memory test. *A* depicts a segment of a recall transcript, with scoring notation in grey. *B* depicts a sample item from the familiarity-judgement task. *C* depicts a sample item from the timeline-judgement task, with scoring depicted in grey. Numbers and "X"s in grey show the true temporal positions of the still frames. Therefore, *i* is the position estimation error for a single still frame, *ii* is the true interval between still frames, *iii* is the judged interval between still frames, and so *iii – ii* is the interval estimation error for this item. *D* depicts a sample item from the comprehension questions.

The recall section consisted of a structured interview with progressive cueing, adapted from the Autobiographical Interview (AI; Levine et al., 2002). The interview began with free recall of the movie, in which the participant was asked to provide as much detail as he/she could remember of the clip. Next, a general probe was given to query for any additional information the participant could think of. Finally, a set of specific probes were used to investigate whether the participant could recall additional information relating to aspects of the settings, characters, and audio-visual details depicted in the clip, as well as the participant's thoughts and feelings experienced while viewing the clip.

Interviews were audio-recorded and transcribed verbatim. For the TLE and HC_{TLE} groups, transcripts were also further anonymized so that raters were blind to group membership. Scoring, based on the AI scoring principles adapted for the movie stimulus, involved segmenting individual details and characterizing them as either internal or external. Internal details were directly related to the content of the movie clip (which could be verified as correct or incorrect) or to the participant's experience of viewing the clip (which were assumed to be correct). External details did not relate directly to the clip (e.g., elaborations based on personal information), were not specific to the clip (e.g., general knowledge), or were repetitions of internal details. The main outcome measure of interest was the total number of correct internal details produced (previously shown to demonstrate adequate inter-rater reliability; see Chapter 3).

In familiarity judgements, participants were shown a single still frame and asked to judge whether it was familiar or unfamiliar. Sixteen items were administered, eight target items (sampled from the entire length of the eight-minute clip shown to participants) and eight lure items (sampled from the unused 12 minutes of the original television episode). A measure of response sensitivity (d′) was calculated based on the rate of hits (rating target items as familiar) and false alarms (rating lures as familiar) in responding.

In timeline judgements, participants were shown two still frames from the movie side by side, as well as a timeline beneath them, and they were asked to click on the timeline where they recalled each still frame to have taken place in the course of the movie. As he/she clicked, an "X" appeared at the selected point on the timeline. The horizontal difference between the leftmost point of the timeline and each mouse click was exported to a datafile and converted into seconds. The task was comprised of 28 items, or 46 still frames, that were sampled from the entire length of the clip with different interval lengths separating them (ranging from 15 to 240 s). A composite measure of performance was computed based on three component measures. First, the number of items in which the two still frames were ordered correctly was tallied. Position estimation error was calculated by subtracting the true temporal position of each still frame from that judged

by the participant, taking the absolute value, and averaging across all still frames. Interval estimation error was calculated by subtracting the true interval separating the two still frames for each item from the interval between the participant's two mouse clicks, taking the absolute value, and averaging across all items. Position and interval estimation error were only measured for trials on which the participant correctly ordered the still frames. The composite measure consisted of the total number of correctly ordered trials (which increases with better performance), divided by the sum of position and interval estimation error (which decrease with better performance).

Finally, twenty comprehension questions were administered used a two-option forced choice format. Some items were accompanied by a still frame from the movie for contextual support. Accuracy was calculated as the total number of correctly answered questions.

5.1.7 Standardized neuropsychological tests

TLE and HC_{TLE} participants completed the following battery of neuropsychological tests. The Names and Doors subtests of the Doors & People Test (Baddeley et al., 2006) were administered to assess verbal and visual recognition memory, respectively. In each task, participants were shown two sets of 12 items (full names or pictures of doors) one at a time, and after each learning phase, their memory was tested using a four-option forced-choice paradigm. The total number of correct responses (out of 24) was used for analysis. The Rey Visual Design Learning Test (RVDLT; Spreen & Strauss, 1991) was included to investigate visual recall memory. Over five learning trials, participants were asked to recall a series of 15 abstract designs, and then recall them again after a 20-minute delay. Delayed recall (out of 15) was used for analysis. The Conditional Associative Learning Test (CALT; Petrides, 1985) was administered as a measure of spatial associative learning with sensitivity to TLE (St-Laurent, McCormick, et al., 2014). For the CALT, participants were asked to learn arbitrary associations between four cards and four randomly placed discs until they achieved 12 consecutive trials correct or completed 68 trials. The number of trials to criterion (between 12 and 68) was used for analysis.

Finally, Matrix Reasoning and Vocabulary – from the Wechsler Adult Intelligence Scale – 4th edition (WAIS-IV; Wechsler, 2008) for the HC_{TLE} group, and from the Wechsler Scale of Abbreviated Intelligence – 2nd edition (WASI-II; Wechsler, 2011) for the TLE group – were included as measures of fluid/nonverbal and crystallized/verbal intelligence, respectively. In Matrix Reasoning, participants are asked to select the missing element of a given matrix from a series of five options. In Vocabulary, participants are asked to provide brief definitions of words. To allow for pooling across the two different versions of Matrix Reasoning and Vocabulary, the age-scaled scores were used for analysis. The HC_{YA} sample also completed the CALT and Matrix Reasoning (WAIS-IV version) as the other groups did.

5.2 Results

5.2.1 Hippocampal volume

The combined right and left hippocampal volumes were significantly reduced in the TLE group compared to the HC_{TLE} group, $t(41) = 3.12$, $p = .003$, $\eta^2 = .19$ (see Figure 11A). Notably, no other subcortical grey matter structures were significantly different across groups. Reduced hippocampal volume was also investigated separately in left and right TLE subgroups (see Figure 11B and Figure 11C). The left TLE group ($n = 8$) showed reduced left hippocampal volume ($t(30) = 2.53$, $p = .017$, $\eta^2 = .18$), but only marginally different right hippocampal volume ($p = .053$, $\eta^2 = .12$) compared to controls. The right TLE group ($n = 9$) showed reduced right hippocampal volume ($t(10) = 2.71$, $p = .022$, $\eta^2 = .19$), but non-significantly different left hippocampal volume ($p = .432$, $\eta^2 = .02$) compared to controls. Therefore, both subgroups demonstrate the expected lateralized patterns of atrophy. Volumetric differences at the individual level are also displayed in Figure 11.

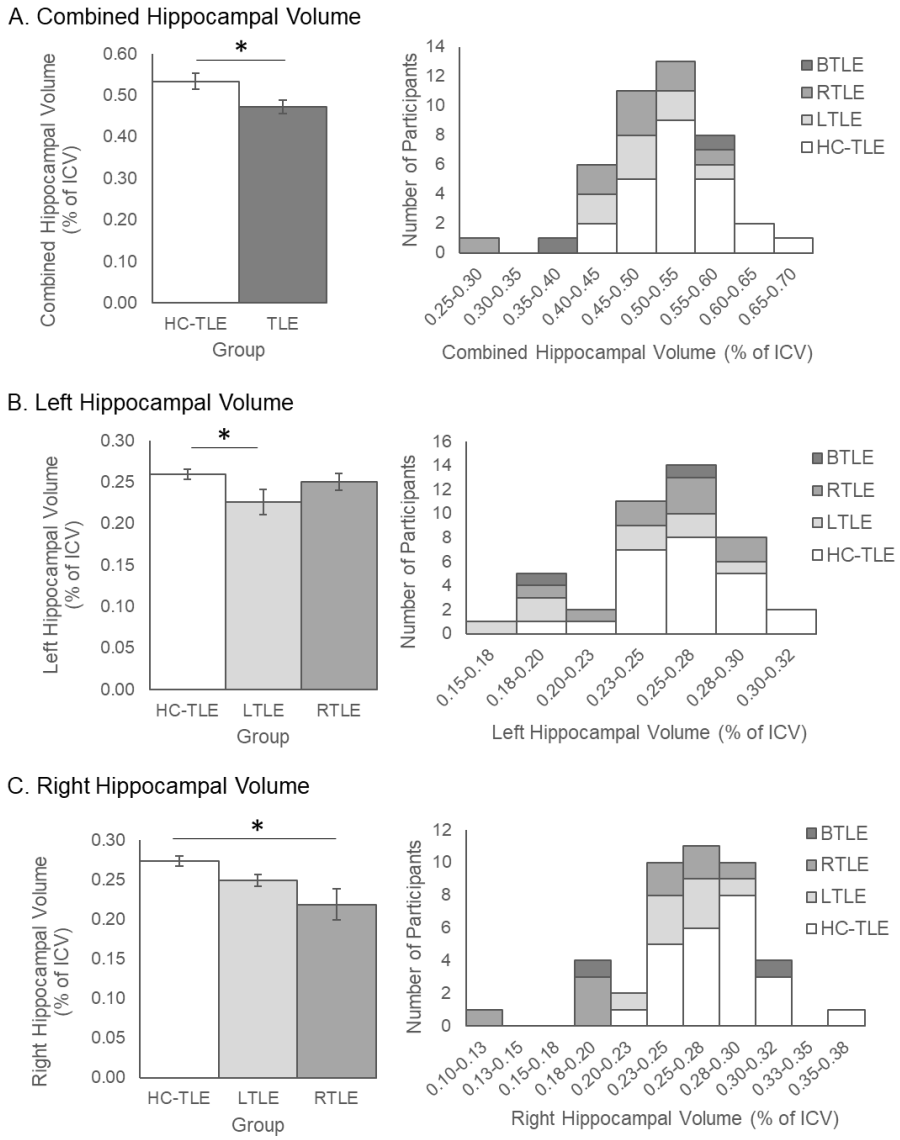


Figure 11. Group comparisons (left) and individual variability (right) in hippocampal volume (expressed as a percentage of the total intracranial volume). *A* depicts the combined left and right hippocampal volumes, *B* depicts the left only, and *C* depicts the right only. Error bars represent standard error of the mean. TLE = temporal lobe epilepsy sample; HC-TLE = healthy control participants – matched to epilepsy sample; BTLE = bilateral TLE; RTLE = right TLE; LTLE = left TLE; ICV = intracranial volume.

* significant at .05 level

The association of hippocampal volume and cognitive performance was first investigated using movie-memory measures. Under FDR correction, three measures of hippocampal volume (left, right, and combined left and right) were correlated with the four memory measures derived from the movie-memory test: total internal details reported in the recall task, sensitivity (d′) on the familiarity-judgement task, a composite measure of performance on the timeline-judgement task, and accuracy on the comprehension questions. In the mixed TLE sample, no correlations between movie-memory measures and combined hippocampal volume or right hippocampal volume survived correction. Left hippocampal volume was correlated with total internal details produced in the recall task (r_s = 0.69, p = .003) and with the composite measure of performance on the timeline-judgement task (r_s = 0.65, p = .007). Pooling the three study samples yielded no significant correlations. To investigate whether null findings in the pooled sample reflected differences in the constituent samples, we statistically compared correlations across the TLE and pooled healthy samples (HC_{YA} and HC_{TLE}) by computing 95% confidence intervals (CI) for the difference between the correlations, where an interval that includes 0 suggests a nonsignificant difference (Zou, 2007). The two significant correlations in the TLE sample for left hippocampal volume described above were significantly stronger in the TLE than healthy sample (95% CI for correlation with recall: [0.38,1.20], and with timeline judgements: [0.18,1.05]), as was the correlation between combined hippocampal volume and recall performance (95% CI: [0.07,1.06]). To summarize, strong correlations between left hippocampal volume and movie-memory measures (from recall and timeline-judgement tasks) in the TLE sample, which were significantly stronger in TLE than in the healthy sample, appear to demonstrate a relationship between hippocampal structural integrity and cognition.

A similar investigation of structure-function relationships was carried out between hippocampal volume and the standardized neuropsychological measures. Correlations between the three measures of hippocampal volume and the six standardized neuropsychological measures were evaluated under FDR correction. In the overall TLE group (both right and left), no correlations survived FDR correction. To improve

statistical power, the three study samples were pooled and the correlations were repeated. Again, no correlations survived FDR correction. Overall, there is no evidence that hippocampal volume was associated with standardized cognitive test performance in the study sample.

5.2.2 Hippocampal connectivity

Whole-brain ROI-to-ROI connectivity was performed across all participants using the hippocampi as seed regions. Numerous regions demonstrate functional coupling with the hippocampi during move viewing (see Appendix E). Notably, the ROI with strongest coupling to the left hippocampus was the right hippocampus. For the right hippocampus, coupling with the left hippocampus was second only to the right posterior parahippocampal gyrus.

5.2.2.1 Interhippocampal coupling

The functional connectivity strength between the hippocampi during movie viewing was quantified in the TLE and HC_{TLE} groups (see Figure 12A). No significant differences in interhippocampal coupling across groups were detected (whole hippocampus: $p = .523$; anterior hippocampus: $p = .505$; posterior hippocampus: $p = .472$). Failure to detect group differences may reflect individual variability in the TLE group (shown in Figure 12B, Figure 12C, and Figure 12D). For each TLE participant, a *z*-score for interhippocampal coupling was computed using the HC_{TLE} group as a normative reference, where a *z* of +/- 1.96 (i.e., 1.96 standard deviations above or below the normative mean) was thought to reflect a significantly discrepant score. Using the whole hippocampal ROI, six of 19 TLE participants had discrepant scores. Of these six, two showed increased hippocampal coupling compared to the normative mean (one nonlesional left TLE, one left TLE with possible left MTS), and four showed decreased coupling (one nonlesional left TLE, two right TLE with right MTS, one nonlesional bilateral TLE). Based on the anterior hippocampal ROIs, seven participants showed discrepant coupling (three increased, four decreased), and based on the posterior hippocampal ROIs, three participants showed discrepant coupling (one increased, two decreased).

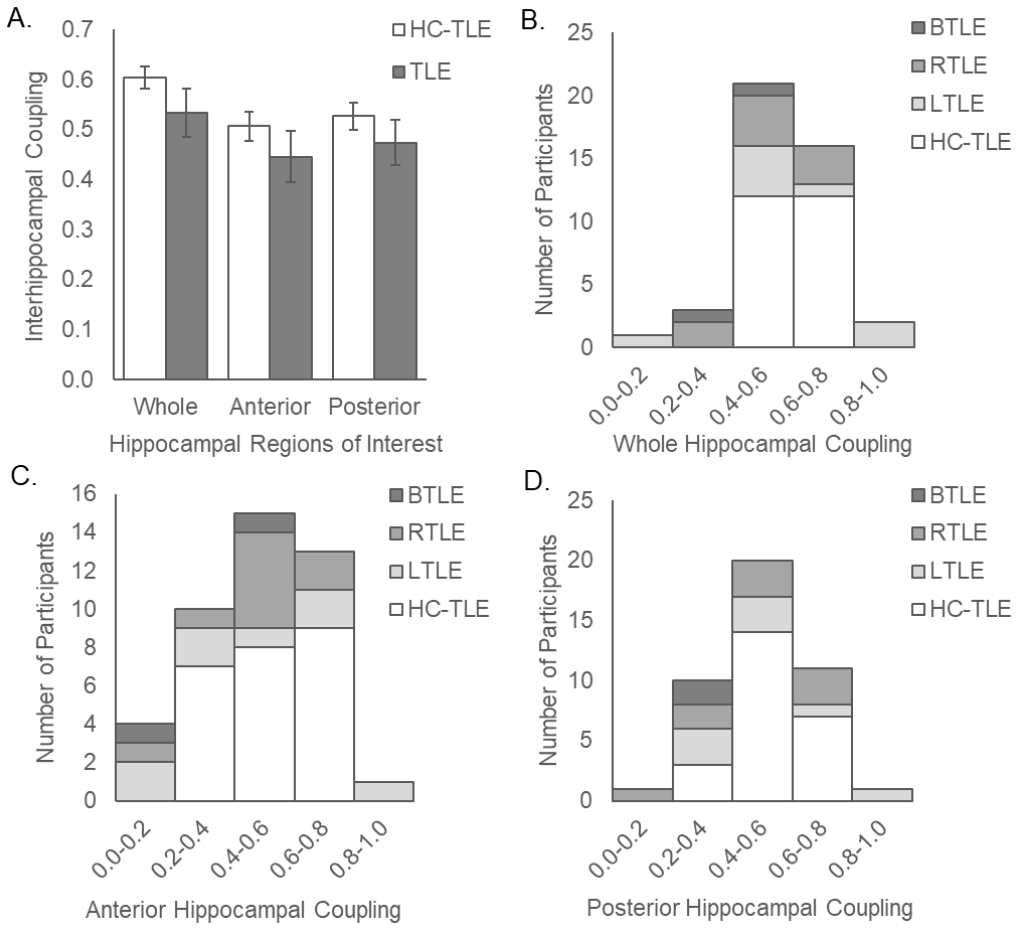


Figure 12. Group comparisons (*A*) and individual variability (*B, C, D*) in interhippocampal coupling. *B* depicts individual variability for whole hippocampal regions-of-interest (ROIs), *C* for anterior hippocampal ROIs, and *D* for posterior hippocampal ROIs. Note that extreme values in the TLE sample occur at both ends of the normative (HC-TLE derived) distribution. Error bars represent standard error of the mean. TLE = temporal lobe epilepsy sample; HC-TLE = healthy control participants – matched to epilepsy sample; BTLE = bilateral TLE; RTLE = right TLE; LTLE = left TLE.

The association of interhippocampal coupling derived from movie-driven fMRI with movie memory was investigated first in the TLE group. Two measures of interhippocampal coupling (based on anterior and posterior hippocampal ROIs) were

correlated with four movie-memory measures under FDR correction. No significant correlations survived correction. Pooling the three study samples yielded two significant correlations. Familiarity-judgement sensitivity significantly correlated with interhippocampal coupling for the anterior hippocampi (r_s = 0.31, p = .013). Additionally, performance on the timeline-judgement task significantly correlated with functional coupling of the anterior hippocampi (r = 0.34, p = .006). No significant correlations emerged for the recall and comprehension tasks. In summary, moderate correlations were detected between functional coupling of the anterior hippocampal ROIs during movie viewing and memory for the movie as tested by the familiarity- and timeline-judgement tasks (see Figure 13), across all tested participants.

Again, to investigate whether non-significant correlations in the pooled sample could be attributed to differences in constituent samples, we statistically compared the correlations between the TLE and combined healthy-control groups. Several correlations were statistically stronger in the TLE sample, including familiarity judgements and whole hippocampal coupling (95% CI: [0.03,0.98]), familiarity judgements and posterior hippocampal coupling (95% CI: [0.03,0.98]), and comprehension and anterior hippocampal coupling (95% CI: [0.01,1.00]). In fact, these three correlations were significant in the TLE sample (p < .05), but they did not survive FDR correction.

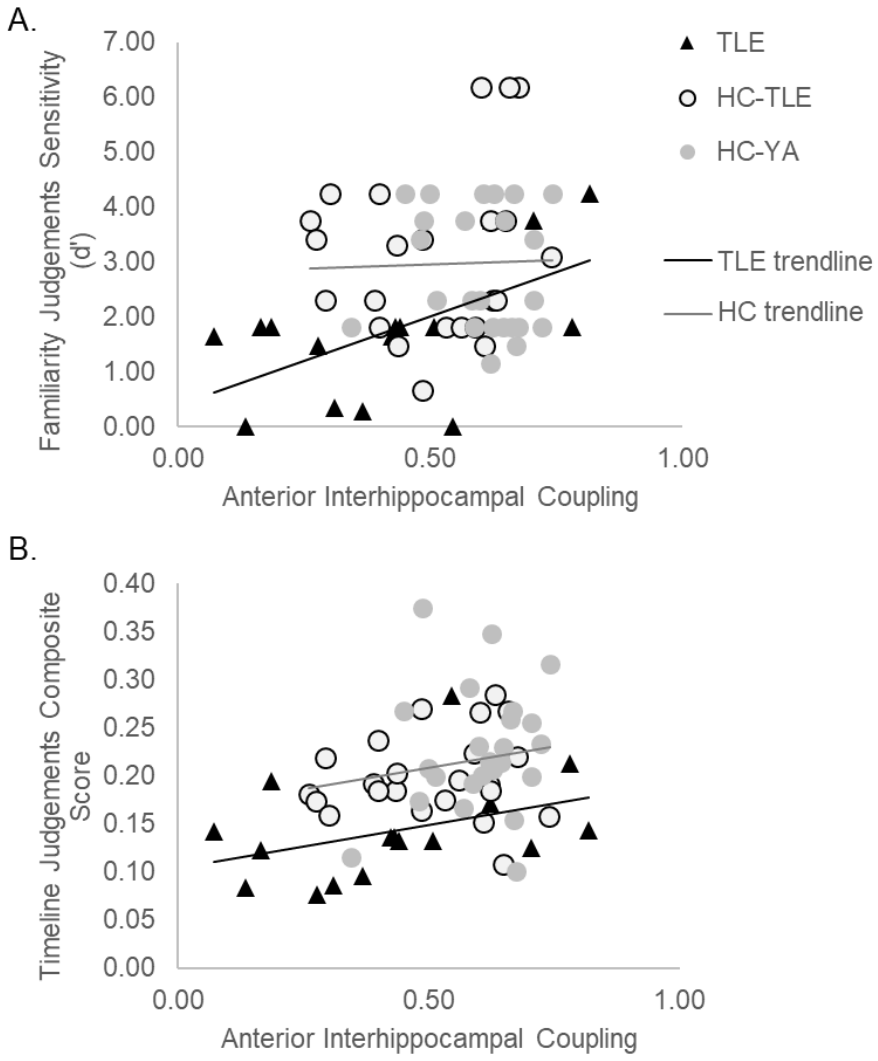


Figure 13. Significant correlations between interhippocampal coupling and movie-memory measures. *A* depicts the moderate positive correlation between familiarity-judgement sensitivity and anterior interhippocampal coupling. *B* depicts the moderate positive correlation between timeline-judgement performance and anterior interhippocampal coupling. Dashed line represents the line of best fit across groups. TLE = temporal lobe epilepsy sample; HC-TLE = healthy control participants – matched to epilepsy sample; HC-YA = healthy control participants – young adults.

For comparison, the association between interhippocampal coupling and standardized neuropsychological measures was also investigated, using the two sets of hippocampal ROIs and six neuropsychological variables under FDR correction. No significant correlations emerged for either the anterior or posterior hippocampal ROIs in the TLE sample alone or when the three study groups were pooled.

5.2.2.2 Hippocampal connectivity with other ROIs

Whole-brain hippocampal connectivity was compared across the TLE and HC_{TLE} groups. Connectivity of the left hippocampus did not significantly differ across groups. However, using the right hippocampus as the source ROI, the medial prefrontal cortex (mPFC) showed significantly weaker connectivity in the TLE compared to the HC_{TLE} group, $t(41) = 4.35$, $p < .001$, $\eta^2 = .32$, which was significant after FDR correction (see Figure 14).

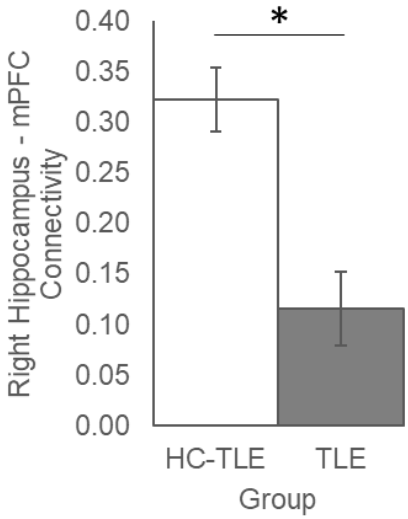


Figure 14. Group differences in functional connectivity of the right hippocampus and medial prefrontal cortex. Error bars represent standard error of the mean. TLE = temporal lobe epilepsy sample; HC-TLE = healthy control participants – matched to epilepsy sample; mPFC = medial prefrontal cortex.

* significant at .05 level

Given this group difference, the association of movie memory performance and functional coupling between the right hippocampus and medial prefrontal cortex during movie viewing was investigated. In the TLE sample, no correlations survived FDR correction. Repeating these correlations in the pooled study sample to increase statistical power, three of the four movie-memory measures significantly correlated with right hippocampus – mPFC connectivity: familiarity-judgement sensitivity (r_s = 0.28, p = .025), comprehension accuracy (r_s = 0.33, p = .009), and timeline-judgement performance (r = 0.35, p = .006). Scatterplots for these correlations are shown in Figure 15. For non-significant correlations in the pooled sample, we again statistically compared the correlations of the constituent samples, and none were significantly different between the TLE and combined healthy-control samples.

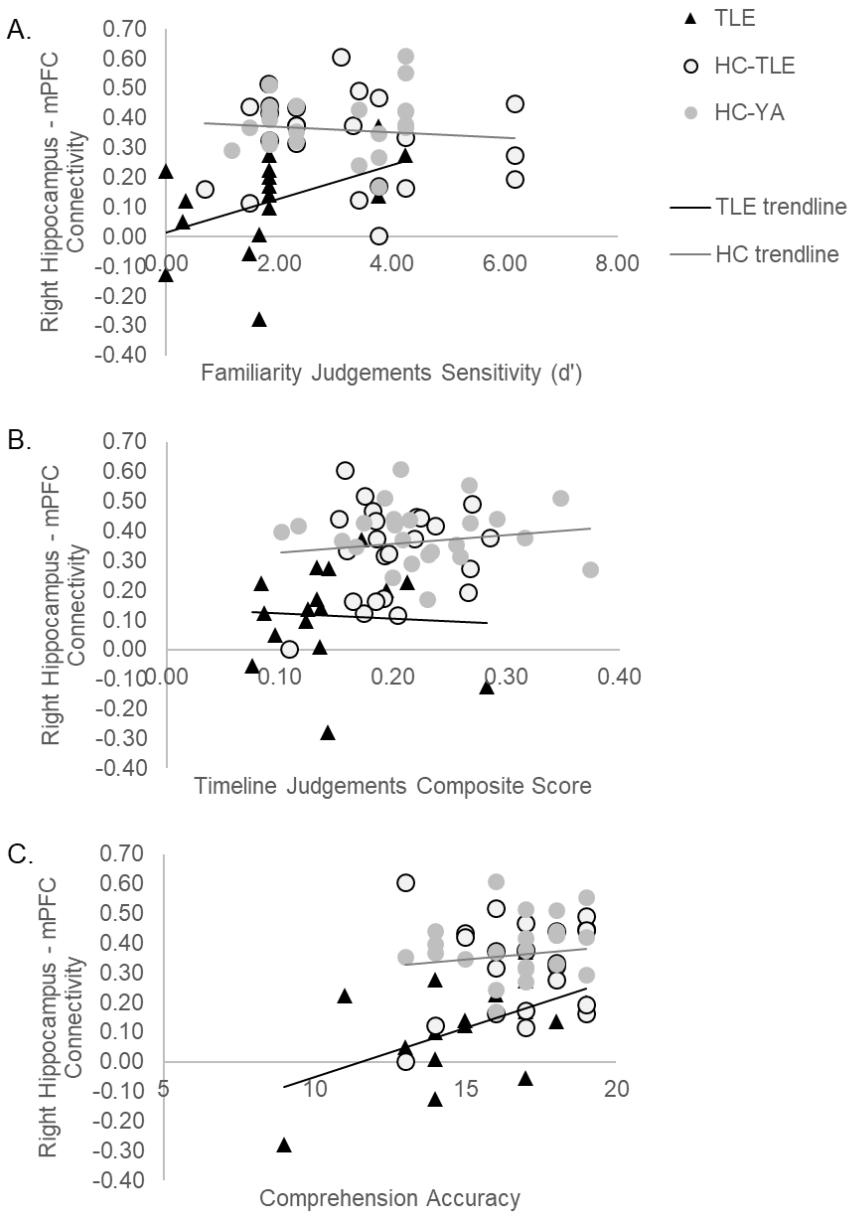


Figure 15. Significant correlations between right hippocampal – medial prefrontal cortex connectivity and movie-memory measures. Moderate positive correlations were detected with outcome measures on the familiarity-judgement task (*A*), the timeline-judgement task (*B*), and the comprehension questions (*C*). Dashed line represents the line of best fit across groups. TLE = temporal lobe epilepsy sample; HC-TLE = healthy control participants – matched to epilepsy sample; HC-YA = healthy control participants – young adults; mPFC = medial prefrontal cortex.

Again, for comparison, the association between right hippocampus – mPFC coupling and standardized neuropsychological measures was also investigated. In the TLE sample, no significant correlations were detected. However, when the study participants were pooled, numerous correlations survive correction, including measures of visual learning and memory (Doors: r_s = 0.47, p = .003; RVDLT: r_s = 0.44, p = .006; CALT: r_s = -0.36, p = .004) and verbal and nonverbal intelligence (Vocabulary: r = 0.44, p = .006; Matrix Reasoning: r = 0.38, p = .003). Thus, movie-driven connectivity between the right hippocampus and mPFC, which was shown to distinguish the two matched samples (TLE and HC_{TLE}), appears to correlate with a range of cognitive measures, both movie-based and standardized.

5.3 Discussion

In the current study, we sought to quantify the relationship between hippocampal volume/connectivity measures and memory abilities in a novel assessment paradigm based on naturalistic stimulation. Functional connectivity of the hippocampus was investigated using movie-driven fMRI, in which participants freely viewed an engaging audiovisual film clip. Memory abilities, in addition to being captured by standardized neuropsychological measures, were also investigated using a multidimensional test of memory for the movie stimulus.

Measures of hippocampal volume based on automated subcortical segmentation reflected the expected patterns of atrophy in TLE participants. In other words, participants with left TLE demonstrated reduced left hippocampal volume, and participants with right TLE demonstrated reduced right hippocampal volume, in keeping with previous literature (Barnett et al., 2015; Bernasconi et al., 2003; Doucet et al., 2016; Fuerst et al., 2001; Lencz et al., 1992; Mechanic-Hamilton et al., 2009). Unlike previous studies (Alessio et al., 2006; Doucet et al., 2016; Lencz et al., 1992; Rausch & Babb, 1993; Reminger et al., 2004; Stoub et al., 2019), we did not detect an association between standardized measures of verbal memory and left or combined hippocampal volume. However, the movie-memory task was sensitive to left hippocampal integrity, and in fact, performance on the

recall and timeline tasks appears to explain, respectively, 48% and 42% of the variability in left hippocampal volume in participants with TLE.

Hippocampal connectivity was investigated in two ways: the strength of functional coupling across the two hippocampi, and whole-brain/hippocampal connectivity differences between the TLE and matched control samples. Altered functional connectivity across homologous medial temporal lobe structures has been documented in TLE, with observations of both increased and decreased connectivity (Maccotta et al., 2013; Morgan et al., 2011; Pittau et al., 2012; Tracy et al., 2014). In the pooled study sample, functional coupling with the contralateral hippocampus was among the strongest functional relationships detected for either hippocampal ROI. Although we did not detect overall group differences in interhippocampal coupling between the TLE and matched control samples, investigation of these coupling strengths at the individual level showed that a number of TLE participants (three to seven of 19 depending on the hippocampal ROI used; whole, anterior, or posterior) demonstrated abnormal coupling, some showing unusually strong coupling compared to controls and others showing unusually weak coupling.

Whether either alteration of interhippocampal coupling (reduced or increased) in TLE represents an index of compromise or effective adaptation remains to be resolved. The present results do not speak to this question, as we did not detect any significant correlations between interhippocampal coupling and memory measures (movie-based or standardized) when correlational analyses were restricted to the TLE sample. In the pooled study sample, however, coupling across anterior hippocampal ROIs was positively correlated with the familiarity and timeline measures of the movie-memory test; that is, stronger coupling was associated with better memory performance. In addition, when we explored differences in the strengths of correlations between TLE and healthy samples, several correlations were significantly stronger in the TLE sample, hinting that TLE has some influence on the concordance of hippocampal coupling and movie memory. We may be underpowered to detect correlations in the TLE sample; this

analysis would ideally be repeated in a larger sample of TLE participants to better understand how the underlying cognitive processes are disrupted or facilitated in the face of altered interhippocampal coupling.

In the whole-brain analysis, the TLE group showed reduced connectivity between the right hippocampus and the mPFC. Furthermore, the strength of this connection was widely correlated with measures of cognition, including positive correlations with performance on both movie-memory (familiarity, timeline, comprehension) and standardized memory (visual recognition memory, visual recall memory, associative memory, crystallized/verbal intelligence, fluid/non-verbal intelligence) measures. Thus, reduced right hippocampus – mPFC connectivity is related to lower cognitive performance.

Both the right hippocampus and mPFC have been shown to support memory for perceptually rich stimuli. St-Laurent and colleagues (2016) found that right hippocampal activation increased during retrieval of perceptually enriched stimuli (film clip) versus perceptually impoverished stimuli (narrative script). As well, right TLE participants showed a dampening of this effect, thought to explain their reduced retrieval of perceptual details. The mPFC is thought to contribute schematic information to episodic memory, relating to the abstract representations of events rather than episode-specific details (Robin & Moscovitch, 2017). Together, the hippocampi and mPFC are nodes of the default-mode network and have been implicated in the retrieval of autobiographical memory (Gilboa, 2004; Spreng, Mar, & Kim, 2009; Svoboda et al., 2006) , with disrupted connectivity in TLE thought to result in impaired autobiographical recall (McCormick et al., 2018).

In the current study, we focused on hippocampal structural integrity and connectivity. To represent the functionally distinct subregions of the hippocampus, we used a coarse bisection of the hippocampus along its anterior-posterior axis. There is support for a functional dissociation between the anterior and posterior hippocampus in episodic memory (Robin & Moscovitch, 2017; Zeidman & Maguire, 2016), that was only

minimally evident in the current study. Other segmentation approaches may yield different results, for example, using functionally rather than anatomically defined ROIs to subsect the hippocampus into anterior and posterior portions (e.g., Voets et al., 2014) or using hippocampal subfields (e.g., dentate gyrus, cornu ammonis regions), which extend along the anterior-posterior axis (e.g., Voets, Hodgetts, Sen, Adcock, & Emir, 2017).

To summarize, we explored brain-behaviour relationships in TLE using a novel assessment paradigm: movie-driven fMRI with subsequent movie-memory testing. Although this is a small sample study, the results are promising: measures of the movie-memory test demonstrated sensitivity to measures of hippocampal integrity and connectivity. Thus, the combination of these movie-based assessment tools offers a direct way of investigating the association between hippocampal network integrity and memory. Converging evidence from complementary neuroimaging and cognitive investigations may improve localization of functionally disrupted and spared networks in TLE, and may ultimately be used to predict and optimize postsurgical outcomes.

6 General discussion

6.1 Contributions

Main findings:

- Movie-based memory measures are sensitive to memory impairment in TLE. Indeed, movie-based measures appear to be more sensitive than several commonly used neuropsychological measures.

- Movie-based memory measures correlated with other memory measures, hippocampal volume, and hippocampal connectivity.

- Movie-derived functional connectivity reveals network alterations that can be directly related to movie memory.

- A list of measures that significantly distinguished the TLE and matched control groups is displayed in Table 15.

General and clinical implications:

- Movie-based tools can be used to assess memory and medial temporal lobe changes in TLE.

- Unlike traditional stimuli, movies provide the opportunity to assess memory for temporal context. Measures of temporal memory can be used to characterize specific episodic memory deficits in TLE, e.g., the reduced influence of event boundaries on the temporal cohesion of memory in TLE.

- Combined movie-driven fMRI and memory testing permits direct integration of different methodologies to investigate brain-behaviour relationships. Converging evidence across companion neuroimaging/cognition tools provides greater assurance regarding the functional integrity of networks for surgical planning.

Table 15. Summary of group comparisons (TLE vs. HC_{TLE}).

Measure Type	$p < .05$	$p > .05$
Standardized cognitive measures	Doors	Names CALT RVDLT
Movie-memory measures	Recall (internal details) Familiarity Judgements (d') Timeline Judgements (composite score) Comprehension (total correct)	
Hippocampal volume	Combined hippocampal volume	
Hippocampal connectivity	Right hippocampus – mPFC connectivity	Interhippocampal coupling Left hippocampus connectivity

Note. CALT = Conditional Associative Learning Test; RVDLT = Rey Visual Design Learning Test; mPFC = medial prefrontal cortex.

Based on a budding literature on the use of movie stimuli in cognitive neuroscience research, we expected that an engaging movie could be used to evoke activity across much of the cortex, reflecting a wide range of processes engaged, and could also capture specific structure-function relationships that capitalize on aspects of the movie's complexity. In this way, movie-based tools may complement the more traditional tests of memory and medial temporal lobe functioning that rely on simple, artificial, unimodal stimuli. To explore the utility of movie-based tools of memory and medial temporal lobe integrity in TLE, we devised a series of investigations (described in Chapters 3 to 5) to test the properties of these tools in healthy and TLE samples, including their concordance with traditional cognitive measures and their sensitivity to TLE. The implications of our findings for the clinical work-up of individuals with refractory TLE are discussed.

6.2 Behavioural measures

Chapters 3 and 4 described the wealth of information about episodic memory that could be derived from the movie-memory test on its own. Chapter 3 focused on the recall,

familiarity, and comprehension sections of the movie-memory test, whereas Chapter 4 focused on the alternative scoring of the recall section (designed to capture temporal aspects of memory) and the timeline-judgement section.

Chapter 3 offered a description of the stages of development of the movie-memory test, an investigation of the sensitivity of movie-memory measures in distinguishing TLE from control participants, and an evaluation of the construct validity of the novel movie-memory measures against standardized cognitive measures. In terms of sensitivity to TLE, measures from each of the test sections – the total number of correct internal details recalled during free and probed recall, the d′ measure of sensitivity in responding on the familiarity-judgement task, and the total number of correctly answered comprehension questions – were shown to be significantly impaired in our sample of TLE participants. Specifically, differences across the TLE and matched control groups explained 13% of the variance in comprehension accuracy, 18% of the variance in familiarity-judgement sensitivity, and 19% of the variance in internal details recalled. When all three of these measures were entered into a regression model to predict group membership, comprehension accuracy was dropped from the model as it did not uniquely contribute to group prediction. Investigating the construct validity of these novel measures in a pooled sample, we demonstrated correlations between our novel task and measures of visual recognition memory, contextualized verbal memory (short stories), and verbal/crystallized intelligence. The significant correlations are also logical on face value. For example, the total number of correct and relevant details recalled for the movie was correlated with the total number of correct details recalled for two short stories (Logical Memory). Similarly, recognition memory for still frames from the movie was correlated with recognition memory for pictures of doors (Doors). Taken together, these results were thought to provide preliminary support for the use of movie-based memory measures to assess medial temporal-lobe dysfunction in refractory TLE.

In Chapter 4, the specific value of temporal-memory measures derived from the movie memory test were explored. Measures of temporal sequencing and of memory for events

at different levels of temporal resolution were computed based on the recall and timeline-judgement sections of the movie-memory test. Just as before, we sought to investigate the sensitivity of these measures in distinguishing TLE from control participants and to explore some basic psychometric properties of the novel measures. Investigating group differences in performance across TLE and matched control participants, all six temporal-memory measures (with the exception of sequencing errors in recall) were significantly impaired in the epilepsy sample. By comparison, only performance on the Doors test (a measure of visual recognition memory) significantly distinguished the groups. Differences due to group accounted for a similar proportion of variance (approximately one third) in performance on the Doors test, in a composite measure of recall performance, and in a composite measure of timeline-judgement performance. In addition, when these measures were entered into logistic regression models, all were found to uniquely contribute to prediction of group (although the recall composite only marginally improved upon prediction based on Doors alone). Using the pooled sample in correlations to investigate construct validity, the recall measures did not significantly correlate with standardized tests but the timeline measures did. Notably, strong correlations were found with tests of visual and verbal memory suggesting that the timeline measures are tapping into an episodic memory construct (or constructs) that includes both visual and verbal encoding/recall modalities. Finally, event-boundary effects on the temporal cohesion of episodic memory were replicated in healthy participants (further evidence of construct validity) and appear to be reduced to some degree in TLE. In summary, the results of this chapter provide support for the disruption of temporal-memory abilities in TLE, and provide preliminary evidence that temporal-memory measures may complement the contributions of other sensitive cognitive measures in assessing cognitive impairment in TLE.

Based on behavioural results described in Chapters 3 and 4, we show promising findings that movie-based memory measures are sensitive to memory dysfunction in individuals with TLE. However, it would be premature to argue for their inclusion in clinical neuropsychological assessments of TLE. Although we demonstrated that temporal-

memory measures could predict TLE status (i.e., classify TLE versus control participants) beyond the predictive value of the Doors test (a strong predictor on its own), this test was just one standardized measure of visual recognition memory. It is uncertain whether, in head-to-head comparison, movie-memory measures would provide unique information to the clinical work-up that would complement standardized neuropsychological measures. One way to investigate this question would be to ask participants (including individuals with TLE and demographically matched controls) to complete both standardized and movie-memory measures (as we did, but perhaps with a broader battery of commonly used standardized measures administered to both groups). With a larger enough sample size to support a regression analysis (conventionally, 15 to 20 participants per regressor), and assuming underlying assumptions were met (e.g., limited multicollinearity), the variables could be entered into a regression model with TLE status as the dependent variable to elucidate the relative predictive values of each measure and the benefit to prediction of combining standardized and movie-memory measures. Beyond TLE status, the use of other dependent variables that are expected to rely on presurgical cognition may also provide evidence for the relative clinical utility of these measures. For example, one of the goals of the neuropsychological assessment is to predict postsurgical memory change. Using memory decline as the dependent variable, one could determine whether movie-memory measures add to the prediction of postsurgical decline.

Our primary motivation in using a movie stimulus was to capitalize on its complexity to capture aspects of cognition that could not be evoked using traditional memory tests due to their simple and artificial stimuli. First, we hoped to provide an adequately rich encoding experience to reveal the reduced vividness in episodic recall of individuals with TLE (St-Laurent, Moscovitch, et al., 2014; St-Laurent et al., 2011). Although individuals with TLE produced fewer details in their recall of the movie, they did not appear to be particularly disadvantaged in producing details that contribute to the vividness of recall. In Chapter 3, they produced fewer story details and fewer perceptual details compared to control participants, with no disproportionate disadvantage for perceptual details. Similarly, in Chapter 4, they produced fewer higher order/indefinite details and fewer

clustered details compared to control participants, with no disproportionate disadvantage for clustered details. We offered several possible explanations as to why a vividness disadvantage was not observed. For example, the movie may provide so many potential story details (disproportionate to the number of perceptual details) that it yields an exaggerated story-detail disadvantage in the TLE group. Further research into the factors that promote or inhibit recall of fine-grained details (e.g., perceptual, clustered) may be valuable in optimizing assessment tools in TLE.

Another aspect of the movie that we harnessed to evaluate memory abilities in TLE was the natural presence of event boundaries that occur in an extended audiovisual narrative. Event boundaries influence the temporal cohesion of episodic memory (DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011, 2014; Faber & Gennari, 2015; Heusser et al., 2018; Zwaan, 1996) and are known to modulate hippocampal activity (Baldassano et al., 2017; Ben-Yakov & Dudai, 2011; Ben-Yakov et al., 2013; Ben-Yakov & Henson, 2018; Clewett et al., 2019; Ezzyat & Davachi, 2011). In Chapter 4, event boundaries were operationalized as scene changes, and a subset of timeline-judgement items (those in which the still frames were sampled 15 or 30 s apart) were specifically designed such that the two still frames were sampled within or across scene changes. Based on previous literature, still frames sampled within a scene were expected to be rated as having occurred closer together in time and were more likely to be ordered correctly than still frames sampled across scene changes. Although the healthy control group showed the expected effect of scene changes on both interval estimation and correct ordering, the TLE group showed the effect only on interval estimation. If the coding of event boundaries relies on hippocampal activation (Ben-Yakov & Henson, 2018) and reduced hippocampal integrity has been associated with aberrant judgements of event boundaries that coincide with impaired sequence memory (Zacks et al., 2006), it follows that individuals with TLE may not experience the same benefit of shared context within the scene to enhance temporal cohesion of moments in the scene. Thus, the temporal complexities of the movie stimulus, like the presence of numerous salient event

boundaries, offer a unique opportunity to investigate the nature of memory deficits in TLE in a way that traditional tests do not.

6.3 Integrating behavioural & neuroimaging measures

In Chapter 5, we explored structural and functional neuroimaging correlates of memory abilities in TLE, capitalizing on the combined movie-driven fMRI/memory protocol. As a measure of structural integrity, hippocampal volumes were computed based on an automated segmentation protocol. TLE participants demonstrated the expected lateralized patterns of atrophy, i.e., on average, those with left TLE showed reduced left hippocampal volume, those with right TLE showed reduced right hippocampal volume, and the contralateral hippocampal volume was in the normal range. Correlations between hippocampal volume and cognitive measures in the TLE sample revealed a significant correlation between the left hippocampal volume and two measures of the movie-memory test (reflecting recall and timeline-judgement performance), with performance on these measures explaining 42 to 48% of the variability in left hippocampal volume in TLE. Turning our attention to functional connectivity of the hippocampus, a number of individual TLE participants showed abnormal interhippocampal connectivity, whether increased or decreased. In the pooled sample, including healthy controls, interhippocampal coupling of the anterior portion of the hippocampus was significantly correlated with two measures from the movie-memory test (familiarity and timeline judgements). In comparing hippocampal connectivity with the whole brain across groups, connectivity of the right hippocampus and medial prefrontal cortex was significantly reduced in TLE, and notably, the strength of this connection was significantly correlated with numerous movie-memory measures, including familiarity judgements, timeline judgements, and comprehension, and with standardized cognitive measures. In summary, the movie-memory test demonstrated sensitivity to measures of hippocampal integrity and connectivity, and abnormal neuroimaging markers could be directly related to cognitive measures.

We set out to investigate whether movie-based assessment tools, including movie-memory testing and movie-driven fMRI could be used to capture episodic memory deficits and medial temporal-lobe dysfunction in TLE. Overall, the movie-memory test was sensitive to cognitive and neural differences in TLE, and the movie-driven fMRI results could be integrated with cognitive outcomes to better understand the functional impact of brain abnormalities. Consider the findings that participants with TLE showed reduced right hippocampus-medial prefrontal cortex connectivity, and reduced connectivity was associated with worse cognitive performance. If we considered either piece of information, the strength of the connection or the performance on cognitive testing, in isolation, we could only make inferences about the nature of their relationship. We might presume that reduced communication across these regions, which are known to be important for episodic memory, might disrupt cognition, just as we might presume that cognitive variables that distinguish a TLE from a control group are reflective of temporal-lobe dysfunction. Directly evaluating the association between these investigations, we can be more confident that they are meaningfully related.

Although it is reassuring to find a relationship between measures of brain function and cognitive abilities, we must also consider whether they are more valuable together than either is alone in detecting temporal-lobe pathology in TLE. Future investigations may replicate and extend the current findings by identifying measures derived from movie-driven fMRI and from movie-memory testing that reliably distinguish TLE participants from healthy participants, and enter these variables into regression models to predict the participant's group (TLE or control). In this way, their relative contributions can be assessed. Of course, this type of analysis would have implications for their ultimate clinical utility as well: if their contributions to detection of temporal-lobe pathology are found to be redundant rather than complementary, then perhaps the more inexpensive option, the memory test, could be used on its own.

6.4 Limitations & conclusion

A major caveat of these investigations is that a number of the analyses, specifically correlational analyses, were conducted in a pooled sample consisting of two healthy control samples (HC_{YA}, HC_{TLE}) and a clinical sample (TLE) to improve statistical power. We cannot assume that the assessment tools measure the same constructs or respond to the same influences in healthy and clinical samples. With a larger sample of TLE participants, we could more confidently explore how specific neural and cognitive changes documented in TLE influence test performance. Furthermore, a larger sample would allow us more clearly to evaluate and distinguish effects of right- and left-medial temporal-lobe dysfunction. Nonetheless, the current findings are promising, and support further investigation of movie-based tools in the presurgical work-up of individuals with refractory TLE.

The investigations also revealed the relative strengths and weaknesses of the movie-memory test's subsections. Statistically, the comprehension section showed the least sensitivity to group and did not uniquely contribute to prediction of group when recall and familiarity-judgement performance had been accounted for. The comprehension questions were included to ensure the participant was following the plot of the movie, and pilot testing revealed that it was relatively easy for healthy young adults to answer these questions correctly. So this result is not surprising, and suggests that the value of the comprehension questions may be as a screening measure to ensure that individuals were paying attention during movie presentation. In terms of practical clinical application, the three computer-administered subsections (familiarity judgements, timeline judgements, and comprehension questions) have an immediate benefit over the orally administered recall interview. Computer administration allows for a fully automated recording and scoring strategy, whereas the recall interview needs to be recorded, transcribed, segmented, and laboriously scored (potentially in two ways) using subjective criteria. This process is not only cumbersome, but it increases measurement error, and complicates comparison against a normative reference sample. It is difficult to evaluate subsections in isolation, since they were always administered in the same order.

The order could not be counterbalanced, since the amount of information provided by the examiner progressively increased with each task. Therefore, future investigations might test subsections individually to better understand their unique contributions to the clinical questions at hand.

Other limitations should be noted. Although we demonstrate a deficit on movie-memory measures in TLE, we cannot be certain that other factors, such as antiepileptic medication use, did not contribute to poor task performance. In addition, the TLE sample was relatively heterogeneous, especially with respect to years since onset (ranging from 2 to 56 years). Those participants with a longer history of epilepsy may be expected to show more substantial cognitive and neural differences, for example, reduced memory and more extensive hippocampal network alterations. In future, these results should be replicated in (a) independent samples of presurgical TLE participants, (b) other TLE samples that were not captured in the present investigations (e.g., drug-free and postsurgical participants), (c) other clinical groups with known temporal-lobe involvement (e.g., dementia, chronic depression), and (d) other epilepsy samples with extratemporal pathology, who would not be expected to show significant impairment. These replications could be used to further investigate the sensitivity of movie-memory measures in correctly identifying individuals with temporal-lobe pathology as well as their specificity in correctly identifying individuals without temporal-lobe pathology.

To conclude, the current investigations provide promising preliminary support for the use of movie-based tools to assess cognitive and neural abnormalities in TLE. Memory testing based on the movie was sensitive to cognitive and neural differences in TLE, and the movie-driven fMRI results could be integrated with movie-memory measures to understand the cognitive implications of functional brain abnormalities. These findings contribute to a growing literature investigating naturalistic viewing of audiovisual stimuli in relation to neuroimaging and memory, with very few studies to date addressing how movie-based measures may be used to understand cognition in clinical populations. As a next step, these tools require further validation, including a demonstration of their unique

contributions above and beyond current gold-standard techniques in predicting temporal-lobe pathology. Beyond that, one could begin to explore the specific clinical questions that may be answered using these novel assessment tools. For example, specific neuroimaging and cognitive markers may be isolated from the breadth of cognitive processes and neural regions recruited during movie viewing to reveal focal areas of impairment at the individual level to assist with seizure localization. At the same time, having functional data pertaining to large networks in the brain may provide valuable information about cognitive resources that can be used to predict postsurgical functioning. More broadly, the application of movie stimuli to assess abnormal brain function was inspired by an evolving body of cognitive-neuroscience research on naturalistic brain processes. Continued translation of cognitive-neuroscience research into clinical-neuropsychological practice ensures that clients benefit from our ever-increasing understanding of brain-behaviour relationships.

References

Abou-Khalil, B. W. (2005). How close is fMRI to providing the memory component of the Wada test? *Epilepsy Currents*, *5*, 184–186.

Addis, D. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus*, *14*, 752–762.

Addis, D. R., Moscovitch, M., & McAndrews, M. P. (2007). Consequences of hippocampal damage across the autobiographical memory network in left temporal lobe epilepsy. *Brain*, *130*, 2327–2342.

Alessio, A., Bonilha, L., Rorden, C., Kobayashi, E., Min, L. L., Damasceno, B. P., & Cendes, F. (2006). Memory and language impairments and their relationships to hippocampal and perirhinal cortex damage in patients with medial temporal lobe epilepsy. *Epilepsy and Behavior*, *8*, 593–600.

Alessio, A., Pereira, F. R. S. S., Sercheli, M. S., Rondina, J. M., Ozelo, H. B., Bilevicius, E., … Cendes, F. (2013). Brain plasticity for verbal and visual memories in patients with mesial temporal lobe epilepsy and hippocampal sclerosis: An fMRI study. *Human Brain Mapping*, *34*, 186–199.

Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (2006). *Doors and People*. Bury St. Edmunds, England: Thames Valley Test Company.

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*, 709–721.

Barnett, A. J., O'Neil, E. B., Watson, H. C., & Lee, A. C. H. (2014). The human hippocampus is sensitive to the durations of events and intervals within a sequence. *Neuropsychologia*, *64*, 1–12.

Barnett, A. J., Park, M. T. M., Pipitone, J., Chakravarty, M. M., & McAndrews, M. P. (2015). Functional and structural correlates of memory in patients with mesial temporal lobe epilepsy. *Frontiers in Neurology*, *6*, 1–9.

Bartels, A., & Zeki, S. (2005). Brain dynamics during natural viewing conditions--a new guide for mapping connectivity in vivo. *NeuroImage*, *24*, 339–349.

Baxendale, S. (2008). The impact of epilepsy surgery on cognition and behavior. *Epilepsy and Behavior*, *12*, 592–599.

Baxendale, S., & Thompson, P. (2010). Beyond localization: The role of traditional neuropsychological tests in an age of imaging. *Epilepsia*, *51*, 2225–2230.

Baxendale, S., Thompson, P., & Duncan, J. (2008). The role of the Wada test in the surgical treatment of temporal lobe epilepsy: An international survey. *Epilepsia*, *49*, 715–720.

Baxendale, S., Thompson, P., Harkness, W., & Duncan, J. (2006). Predicting memory decline following epilepsy surgery: A multivariate approach. *Epilepsia*, *47*, 1887–1894.

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magentic resonance imaging. *IEEE Transactions on Medical Imaging*, *23*, 137–152.

Bell, B. D., & Davies, K. G. (1998). Anterior temporal lobectomy, hippocampal sclerosis, and memory: Recent neuropsychological findings. *Neuropsychology Review*, *8*, 25–41.

Ben-Yakov, A., & Dudai, Y. (2011). Constructing realistic engrams: Poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *Journal of Neuroscience, 31*, 9032–9042.

Ben-Yakov, A., Eshel, N., & Dudai, Y. (2013). Hippocampal immediate poststimulus activity in the encoding of consecutive naturalistic episodes. *Journal of Experimental Psychology: General, 142*, 1255–1263.

Ben-Yakov, A., & Henson, R. N. (2018). The hippocampal film editor: Sensitivity and specificity to event boundaries in continuous experience. *Journal of Neuroscience*, *38*, 10057–10068.

Ben-Yakov, A., Rubinson, M., & Dudai, Y. (2014). Shifting gears in hippocampus: Temporal dissociation between familiarity and novelty signatures in a single event. *Journal of Neuroscience, 34*, 12973–12981.

Benjamin, C. F. A., Li, A. X., Blumenfeld, H., Constable, R. T., Alkawadri, R., Bickel, S., … Hirsch, L. J. (2018). Presurgical language fMRI: Clinical practices and patient outcomes in epilepsy surgical planning. *Human Brain Mapping, 39*, 2777–2785.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological), 57*, 289–300.

Berkovic, S. F., Andermann, F., Olivier, A., Ethier, R., Melanson, D., Robitaille, Y., … Feindel, W. (1991). Hippocampal sclerosis in temporal lobe epilepsy demonstrated by magnetic resonance imaging. *Annals of Neurology, 29*, 175–182.

Bernasconi, N., Bernasconi, A., Caramanos, Z., Antel, S. B., Andermann, F., & Arnold, D. L. (2003). Mesial temporal damage in temporal lobe epilepsy: A volumetric MRI study of the hippocampus, amygdala and parahippocampal region. *Brain, 126*, 462–469.

Betti, V., DellaPenna, S., de Pasquale, F., Mantini, D., Marzetti, L., Romani, G. L., & Corbetta, M. (2013). Natural scenes viewing alters the dynamics of functional connectivity in the human brain. *Neuron, 79*, 782–797.

Bettus, G., Guedj, E., Joyeux, F., Confort-Gouny, S., Soulier, E., Laguitton, V., … Guye, M. (2009). Decreased basal fMRI functional connectivity in epileptogenic networks

and contralateral compensatory mechanisms. *Human Brain Mapping*, *30*, 1580–1591.

Binder, J. R. (2012). Preoperative prediction of verbal episodic memory outcome using fMRI. *Neurosurgery Clinics of North America*, *22*, 219–232.

Binder, J. R., Sabsevitz, D. S., Swanson, S. J., Hammeke, T. A., Raghavan, M., & Mueller, W. M. (2008). Use of preoperative functional MRI to predict verbal memory decline after temporal lobe epilepsy surgery. *Epilepsia*, *49*, 1377–1394.

Bird, C. M., & Burgess, N. (2008). The hippocampus and memory: Insights from spatial processing. *Nature Reviews Neuroscience*, *9*, 182–194.

Bonnici, H. M., Sidhu, M., Chadwick, M. J., Duncan, J. S., & Maguire, E. A. (2013). Assessing hippocampal functional reserve in temporal lobe epilepsy: A multi-voxel pattern analysis of fMRI data. *Epilepsy Research*, *105*, 140–149.

Bowles, B., Crupi, C., Mirsattari, S. M., Pigott, S. E., Parrent, A. G., Pruessner, J. C., … Kohler, S. (2007). Impaired familiarity with preserved recollection after anterior temporal-lobe resection that spares the hippocampus. *Proceedings of the National Academy of Sciences*, *104*, 16382–16387.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Branch, C., Milner, B., & Rasmussen, T. (1964). Intracarotid sodium amytal for the lateralization of cerebral speech dominance: Observations in 123 patients. *Journal of Neurosurgery*, *21*, 399–405.

Briellmann, R. S., Kalnins, R. M., Berkovic, S. F., & Jackson, G. D. (2002). Hippocampal pathology in refractory temporal lobe epilepsy: T2-weighted signal change reflects dentate gliosis. *Neurology*, *58*, 265–271.

Brunec, I. K., Ozubko, J. D., Barense, M. D., & Moscovitch, M. (2017). Recollection-dependent memory for event duration in large-scale spatial navigation. *Learning & Memory*, *24*, 104–114.

Burgess, N., Becker, S., King, J. A., & O'Keefe, J. (2001). Memory for events and their spatial context: Models and experiments. *Phil Trans R Soc Lond B*, *356*, 1493–1503.

Campbell, D., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*.

Campbell, K. L., Shafto, M. A., Wright, P., Tsvetanov, K. A., Geerligs, L., Cusack, R., … Tyler. (2015). Idiosyncratic responding during movie-watching predicted by age differences in attentional control. *Neurobiology of Aging*, *36*, 3045–3055.

Carne, R. P., O'Brien, T. J., Kilpatrick, C. J., MacGregor, L. R., Hicks, R. J., Murphy, M. A., … Cook, M. J. (2004). MRI-negative PET-positive temporal lobe epilepsy: A distinct surgically remediable syndrome. *Brain*, *127*, 2276—2285.

Cataldi, M., Avoli, M., & De Villers-Sidani, E. (2013). Resting state networks in temporal lobe epilepsy. *Epilepsia*, *54*, 2048–2059.

Centeno, M., Tierney, T. M., Perani, S., Shamshiri, E. A., StPier, K., Wilkinson, C., … Carmichael, D. W. (2016). Optimising EEG-fMRI for localisation of focal epilepsy in children. *PLoS ONE*, *11*, e0149048.

Chelune, G. J. (1995). Hippocampal adequacy versus functional reserve: Predicting memory functions following temporal lobectomy. *Archives of Clinical Neuropsychology*, *10*, 413–432.

Cherbuin, N., Anstey, K. J., Réglade-Meslin, C., & Sachdev, P. S. (2009). In vivo hippocampal measurement and memory: A comparison of manual tracing and automated segmentation in a large community-based sample. *PLoS ONE*, *4*, e5265.

Chiba, A. A., Kesner, R. P., & Reynolds, A. M. (1994). Memory for spatial location as a function of temporal lag in rats: Role of hippocampus and medial prefrontal cortex. *Behavioral and Neural Biology*, *61*, 123–131.

Chun, M. M., & Phelps, E. A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience*, *2*, 844–847.

Cicchetti, D. V. (1994). Guidlines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290.

Clewett, D., Dubrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*, 162–183.

Concha, L., Beaulieu, C., Collins, D. L., & Gross, D. W. (2009). White-matter diffusion abnormalities in temporal-lobe epilepsy with and without mesial temporal sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *80*, 312–319.

Daselaar, S. M., Rice, H. J., Greenberg, D. L., Cabeza, R., LaBar, K. S., & Rubin, D. C. (2008). The spatiotemporal dynamics of autobiographical memory: Neural correlates of recall, emotional intensity, and reliving. *Cerebral Cortex*, *18*, 217–229.

Datta, A., & Loddenkemper, T. (2011). The epileptogenic zone. In E. Wyllie (Ed.), *Treatment of Epilepsy: Principles and Practice* (Fifth, pp. 818–827). Philadelphia: Wolters Kluwer.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*, 693–700.

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, *19*, 92–99.

de Campos, B. M., Coan, A. C., Lin Yasuda, C., Casseb, R. F., & Cendes, F. (2016). Large-scale brain networks are distinctly affected in right and left mesial temporal lobe epilepsy. *Human Brain Mapping*, *37*, 3137–3152.

Dede, A. J. O., Frascino, J. C., Wixted, J. T., & Squire, L. R. (2016). Learning and remembering real-world events after medial temporal lobe damage. *Proceedings of*

the National Academy of Sciences, *113*, 13480–13485.

Del Felice, A., Beghi, E., Boero, G., La Neve, A., Bogliun, G., De Palo, A., & Specchio, L. M. (2010). Early versus late remission in a cohort of patients with newly diagnosed epilepsy. *Epilepsia*, *51*, 37–42.

Delis, D. C., Kaplan, E., Kramer, J., & Ober, B. (2000). *California Verbal Learning Test* (2nd ed.). San Antonio: Pearson.

Detre, J. A., Maccotta, L., King, D., Alsop, D. C., Glosser, G., D'Esposito, M., … French, J. A. (1998). Functional MRI lateralization of memory in temporal lobe epilepsy. *Neurology*, *50*, 926–932.

Deuker, L., Bellmund, J. L., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *ELife*, *5*, 1–26.

Diniz, P. R. B., Velasco, T. R., Salmon, C. E. G., Sakamoto, A. C., Leite, J. P., & Santos, A. C. (2011). Extratemporal damage in temporal lobe epilepsy: Magnetization transfer adds information to volumetric MR imaging. *American Journal of Neuroradiology*, *32*, 1857–1861.

Doucet, G. E., He, X., Sperling, M., Sharan, A., & Tracy, J. I. (2016). Gray matter abnormalities in temporal lobe epilepsy: Relationships with resting-state functional connectivity and episodic memory performance. *PLoS ONE*, *11*, 1–21.

Doucet, G. E., Osipowicz, K., Sharan, A., Sperling, M. R., & Tracy, J. I. (2013). Extratemporal functional connectivity impairments at rest are related to memory performance in mesial temporal epilepsy. *Human Brain Mapping*, *34*, 2202–2216.

Dubrow, S., & Davachi, L. (2014). Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. *Journal of Neuroscience*, *34*, 13998–14005.

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *J Exp Psychol Gen*, *142*, 1277–1286.

Duchowny, M., Jayakar, P., Harvey, A. S., Resnick, T., Alvarez, L., Dean, P., & Levin, B. (1996). Language cortex representation: Effects of developmental versus acquired pathology. *Annals of Neurology*, *40*, 31–38.

Dudukovic, N. M., & Wagner, A. D. (2007). Goal-dependent modulation of declarative memory: Neural correlates of temporal recency decisions and novelty detection. *Neuropsychologia*, *45*, 2608–2620.

Eichenbaum, H. (2000). A cortical–hippocampal system for declarative memory. *Nature Reviews Neuroscience*, *1*, 41–50.

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*, 732–744.

Ekstrom, A. D., & Ranganath, C. (2018). Space, time, and episodic memory: The hippocampus is all over the cognitive map. *Hippocampus*, *28*, 680–687.

Engel, J. (1996). Surgery for seizures. *New England Journal of Medicine*, *334*, 647–652.

Engel, J. (2001). Mesial temporal lobe epilepsy: What have we learned? *The Neuroscientist*, *7*, 340–352.

Engel, J., Thompson, P. M., Stern, J. M., Staba, R. J., Bragin, A., & Mody, I. (2013). Connectomics and epilepsy. *Current Opinion in Neurology*, *26*, 186–194.

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychological Science*, *22*, 243–252.

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*, 1179–1189.

Faber, M., & Gennari, S. P. (2015). In search of lost time: Reconstructing the unfolding of events from memory. *Cognition*, *143*, 193–202.

Falconer, M. A., Serafetinides, E. A., & Corsellis, J. A. N. (1964). Etiology and pathogenesis of temporal lobe epilepsy. *Archives of Neurology*, *10*, 233–348.

Fisher, R. S., Cross, J. H., French, J. A., Higurashi, N., Hirsch, E., Jansen, F. E., … Zuberi, S. M. (2017). Operational classification of seizure types by the International League Against Epilepsy: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*, *58*, 522–530.

Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, 458–462.

Fountas, K. N. (2011). Implanted subdural electrodes: Safety issues and complication avoidance. *Neurosurgery Clinics of North America*, *22*, 519–531.

Fuerst, D., Shah, J., Kupsky, W. J., Johnson, R., Shah, A., Hayman-Abello, B., … Watson, C. (2001). Volumetric MRI, pathological, and neuropsychological progression in hippocampal sclerosis. *Neurology*, *57*, 184–188.

Furman, O., Dorfman, N., Hasson, U., Davachi, L., & Dudai, Y. (2007). They saw a movie: Long-term memory for an extended audiovisual narrative. *Learning & Memory*, *14*, 457–467.

Furman, O., Mendelsohn, A., & Dudai, Y. (2012). The episodic engram transformed: Time reduces retrieval-related brain activity but correlates it with memory accuracy. *Learning and Memory*, *19*, 575–587.

Geerligs, L., Rubinov, M., Cam-CAN, & Henson, R. N. (2015). State and trait components of functional connectivity: Individual differences vary with mental state. *Journal of Neuroscience*, *35*, 13949–13961.

Gilboa, A. (2004). Autobiographical and episodic memory - one and the same? Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia*, *42*, 1336–1349.

Gleissner, U., Helmstaedter, C., & Elger, C. E. (1998). Right hippocampal contribution to visual memory: A presurgical and postsurgical study in patients with temporal lobe

epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry*, *65*, 665–669.

Gleissner, U., Helmstaedter, C., & Elger, C. E. (2002). Memory reorganization in adult brain: Observations in three patients with temporal lobe epilepsy. *Epilepsy Research*, *48*, 229–234.

Glikmann-Johnston, Y., Saling, M. M., Chen, J., Cooper, K. A., Beare, R. J., & Reutens, D. C. (2008). Structural and functional correlates of unilateral mesial temporal lobe spatial memory impairment. *Brain*, *131*, 3006–3018.

Glosser, G., Saykin, A. J., Deutsch, G. K., O'Connor, M. J., & Sperling, M. R. (1995). Neural Organization of Material-Specific Memory Functions in Temporal Lobe Epilepsy Patients as Assessed by the Intracarotid Amobarbital Test. *Neuropsychology*, *9*, 449–456.

Golby, A. J., Poldrack, R. A., Illes, J., Chen, D., Desmond, J. E., & Gabrieli, J. D. E. (2002). Memory lateralization in medial temporal lobe epilepsy assessed by functional MRI. *Epilepsia*, *43*, 855–863.

Griffanti, L., Douaud, G., Bijsterbosh, J., Evangelisti, S., Alfaro-Almagro, F., Glasser, M. F., … Smith, S. M. (2017). Hand classification of fMRI ICA noise components. *NeuroImage*, *154*, 188–205.

Griffanti, L., Salimi-Khorshidi, G., Beckmann, C. F., Auerbach, E. J., Douaud, G., Sexton, C. E., … Smith, S. M. (2014). ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage*, *95*, 232–247.

Hales, J. B., & Brewer, J. B. (2010). Activity in the hippocampus and neocortical working memory regions predicts successful associative memory for temporally discontiguous events. *Neuropsychologia*, *48*, 3351–3359.

Hales, J. B., & Brewer, J. B. (2011). The timing of associative memory formation: frontal lobe and anterior medial temporal lobe activity at associative binding predicts memory. *Journal of Neurophysiology*, *105*, 1454–1463.

Hamberger, M. J., & Cole, J. (2011). Language organization and reorganization in epilepsy. *Neuropsychology Review*, *21*, 240–251.

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi, L. (2008). Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron*, *57*, 452–462.

Hasson, U., & Honey, C. J. (2012). Future trends in Neuroimaging: Neural processes as expressed within real-life contexts. *NeuroImage*, *62*, 1272–1278.

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

Hasson, U., & Malach, R. (2008). Human brain activation during viewing of dynamic natural scenes. In *Percept, Decision, Action: Bridging the Gaps* (pp. 203–216).

Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, *14*, 40–48.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*, 1634–1640.

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*, 2539–2550.

Helmstaedter, C., Kurthen, M., Lux, S., Reuber, M., & Elger, C. E. (2003). Chronic epilepsy and cognition: A longitudinal study in temporal lobe epilepsy. *Annals of Neurology*, *54*, 425–432.

Helmstaedter, C., Pohl, C., Hufnagel, A., & Elger, C. E. (1991). Visual learning deficits in nonresected patients with right temporal lobe epilepsy. *Cortex*, *27*, 547–555.

Herfurth, K., Kasper, B., Schwarz, M., Stefan, H., & Pauli, E. (2010). Autobiographical memory in temporal lobe epilepsy: Role of hippocampal and temporal lateral structures. *Epilepsy and Behavior*, *19*, 365–371.

Hermann, B. P., Seidenberg, M., Haltiner, A., & Wyler, A. R. (1995). Relationship of age at onset, chronologic age, and adequacy of preoperative performance to verbal memory change after anterior temporal lobectomy. *Epilepsia*, *36*, 137–145.

Hermann, B. P., Seidenberg, M., Schoenfeld, J., & Davies, K. (1997). Neuropsychological characteristics of the syndrome of mesial temporal lobe epilepsy. *Archives of Neurology*, *54*, 369–376.

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

Holmes, M., Folley, B. S., Sonmezturk, H. H., Gore, J. C., Kang, H., Abou-Khalil, B., & Morgan, V. L. (2014). Resting state functional connectivity of the hippocampus associated with neurocognitive function in left temporal lobe epilepsy. *Human Brain Mapping*, *35*, 735–744.

Hong, K. S., Lee, S. K., Kim, J. Y., Lee, D. S., & Chung, C. K. (2002). Pre-surgical evaluation and surgical outcome of 41 patients with non-lesional neocortical epilepsy. *Seizure*, *11*, 184—192.

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General*, *142*, 1211–1230.

Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of Neuroscience*, *34*, 4692–4707.

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, *81*, 1165–1178.

Huijbers, W., Van Dijk, K. R. A., Boenniger, M. M., Stirnberg, R., & Breteler, M. M. B. (2017). Less head motion during MRI under task than resting-state conditions. *NeuroImage*, *147*, 111–120.

Jacques, P. S., Rubin, D. C., Labar, K. S., & Cabeza, R. (2008). The short and long of it: Neural correlates of temporal-order memory for autobiographical events. *Journal of Cognitive Neuroscience*, *20*, 1327–1341.

James, G. A., Tripathi, S. P., Ojemann, J. G., Gross, R. E., & Drane, D. L. (2013). Diminished default mode network recruitment of the hippocampus and parahippocampus in temporal lobe epilepsy. *Journal of Neurosurgery*, *119*, 288–300.

Janecek, J. K., Swanson, S. J., Sabsevitz, D. S., Hammeke, T. A., Raghavan, M., E. Rozman, M., & Binder, J. R. (2013). Language lateralization by fMRI and Wada testing in 229 patients with epilepsy: Rates and predictors of discordance. *Epilepsia*, *54*, 314–322.

Jenkins, L. J., & Ranganath, C. (2010). Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *Journal of Neuroscience*, *30*, 15558–15565.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*, 825–841.

Jones-Gotman, M., Smith, M. Lou, Risse, G. L., Westerveld, M., Swanson, S. J., Giovagnoli, A. R., … Piazzini, A. (2010). The contribution of neuropsychology to diagnostic assessment in epilepsy. *Epilepsy and Behavior*, *18*, 3–12.

Kennepohl, S., Sziklas, V., Garver, K. E., Wagner, D. D., & Jones-Gotman, M. (2007). Memory and the medial temporal lobe: Hemispheric specialization reconsidered. *NeuroImage*, *36*, 969–978.

Kesner, R. P., Gilbert, P. E., & Barua, L. A. (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. *Behavioral Neuroscience*, *116*, 286–290.

Kim, H., Yi, S., Son, E. I., & Kim, J. (2003). Material-specific memory in temporal lobe epilepsy: Effects of seizure laterality and language dominance. *Neuropsychology*, *17*, 59–68.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, ECVP Abstract Supplement.

Konishi, S., Asari, T., Jimura, K., Chikazoe, J., & Miyashita, Y. (2006). Activation shift from medial to lateral temporal cortex associated with recency judgements following

impoverished encoding. *Cerebral Cortex*, *16*, 469–474.

Konkel, A. (2008). Hippocampal amnesia impairs all manner of relational memory. *Frontiers in Human Neuroscience*, *2*, 1–15.

Kuba, R., Tyrlíková, I., Chrastina, J., Slaná, B., Pažourková, M., Hemza, J., … Rektor, I. (2011). "MRI-negative PET-positive" temporal lobe epilepsy: Invasive EEG findings, histopathology, and postoperative outcomes. *Epilepsy and Behavior*, *22*, 537–541.

Kumaran, D., & Maguire, E. A. (2009). Novelty signals: A window into hippocampal information processing. *Trends in Cognitive Sciences*, *13*, 47–54.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*, 72–79.

Kwan, P., & Brodie, M. J. (2000). Early identification of refractory epilepsy. *New England Journal of Medicine*, *342*, 314–319.

Kwan, P., & Brodie, M. J. (2010). Definition of refractory epilepsy: Defining the indefinable? *Lancet Neurology*, *9*, 27–29.

Lee, T. M. C., Yip, J. T. H., & Jones-Gotman, M. (2002). Memory deficits after resection from left or right anterior temporal lobe in humans: A meta-analytic review. *Epilepsia*, *43*, 283–291.

Lehn, H., Steffenach, H.-A., van Strien, N. M., Veltman, D. J., Witter, M. P., & Haberg, A. K. (2009). A specific role of the human hippocampus in recall of temporal sequences. *Journal of Neuroscience*, *29*, 3475–3484.

Lencz, T., McCarthy, G., Bronen, R. A., Scott, T. M., Inserni, J. A., Sass, K. J., … Spencer, D. D. (1992). Quantitative magnetic resonance imaging in temporal lobe epilepsy: Relationship to neuropathology and neuropsychological function. *Annals of Neurology*, *31*, 629–637.

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, *17*, 677–689.

Lewis, D. V., Barboriak, D. P., MacFall, J. R., Provenzale, J. M., Mitchell, T. V., & VanLandingham, K. E. (2002). Do prolonged febrile seizures produce medial temporal sclerosis? Hypotheses, MRI evidence and unanswered questions. In *Progress in Brain Research* (pp. 263–278).

Liao, W., Zhang, Z., Pan, Z., Mantini, D., Ding, J., Duan, X., … Chen, H. (2011). Default mode network abnormalities in mesial temporal lobe epilepsy: A study combining fMRI and DTI. *Human Brain Mapping*, *32*, 883–895.

Limotai, C., & Mirsattari, S. M. (2012). Role of functional MRI in presurgical evaluation of memory function in temporal lobe epilepsy. *Epilepsy Research and Treatment*, 1–12.

Liu, M., Chen, Z., Beaulieu, C., & Gross, D. W. (2014). Disrupted anatomic white matter network in left mesial temporal lobe epilepsy. *Epilepsia*, *55*, 674–682.

Loddenkemper, T., & Staudt, M. (2011). Eloquent Cortex and the Role of Plasticity. In E. Wyllie (Ed.), *Treatment of Epilepsy: Principles and Practice* (Fifth, pp. 887–898). Philadelphia: Wolters Kluwer.

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Poppenk, J. L., Hasson, U., & Norman, K. A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *ELife*, *5*, e16070.

Maccotta, L., He, B. J., Snyder, A. Z., Eisenman, L. N., Benzinger, T. L., Ances, B. M., … Hogan, R. E. (2013). Impaired and facilitated functional networks in temporal lobe epilepsy. *NeuroImage: Clinical*, *2*, 862–872.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, *71*, 737–749.

Maguire, E. A. (2012). Studying the freely-behaving brain with fMRI. *NeuroImage*, *62*, 1170–1176.

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*, 530–540.

Mattson, R. H., Cramer, J. A., & Collins, J. F. (1996). Prognosis for total control of complex partial and secondarily generalized tonic clonic seizures. *Neurology*, *47*, 68–76.

Mayes, A., Montaldi, D., & Migo, E. (2007). Associative memory and the medial temporal lobes. *Trends in Cognitive Sciences*, *11*, 126–135.

McAndrews, M. P. (2014). Memory assessment in the clinical context using functional magnetic resonance imaging. *Neuroimaging Clinics of North America*, *24*, 585–597.

McAndrews, M. P., & Cohn, M. (2012). Neuropsychology in temporal lobe epilepsy: Influences from cognitive neuroscience and functional neuroimaging. *Epilepsy Research and Treatment*, 1–13.

McCormick, C., Moscovitch, M., Valiante, T. A., Cohn, M., & McAndrews, M. P. (2018). Different neural routes to autobiographical memory recall in healthy people and individuals with left medial temporal lobe epilepsy. *Neuropsychologia*, *110*, 26–36.

McCormick, C., Protzner, A. B., Barnett, A. J., Cohn, M., Valiante, T. A., & McAndrews, M. P. (2014). Linking DMN connectivity to episodic memory capacity: What can we learn from patients with medial temporal lobe damage? *NeuroImage: Clinical*, *5*, 188–196.

McCormick, C., Quraan, M., Cohn, M., Valiante, T. A., & McAndrews, M. P. (2013).

Default mode network connectivity indicates episodic memory capacity in mesial temporal lobe epilepsy. *Epilepsia*, *54*, 809–818.

McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia*, *47*, 2290–2298.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30–46.

Mechanic-Hamilton, D., Korczykowski, M., Yushkevich, P. A., Lawler, K., Pluta, J., Glynn, S., … Detre, J. A. (2009). Hippocampal volumetry and functional MRI of memory in temporal lobe epilepsy. *Epilepsy and Behavior*, *16*, 128–138.

Milner, B. (1970). Memory and the medial temporal regions of the brain. In K. H. Pribram & D. E. Broadbent (Eds.), *Biology of Memory*. New York: Academic Press.

Milner, B. (1972). Disorders of learning and memory after temporal lobe lesions in man. *Clinical Neurosurgery*, *19*, 421–446.

Möddel, G., Lineweaver, T., Schuele, S. U., Reinholz, J., & Loddenkemper, T. (2009). Atypical language lateralization in epilepsy patients. *Epilepsia*, *50*, 1505–1516.

Mohanraj, R., & Brodie, M. J. (2006). Determining pharmaceutical intractibility. In J. W. Miller & D. L. Silbergeld (Eds.), *Epilepsy Surgery: Principles and Controversies* (pp. 3–19). New York: Taylor & Francis.

Montchal, M. E., Reagh, Z. M., & Yassa, M. A. (2019). Precise temporal memories are supported by the lateral entorhinal cortex in humans. *Nature Neuroscience*, *22*, 284–288.

Morey, R. A., Petty, C. M., Xu, Y., Pannu Hayes, J., Wagner, H. R., Lewis, D. V., … McCarthy, G. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *NeuroImage*, *45*, 855–866.

Morgan, V. L., Rogers, B. P., Sonmezturk, H. H., Gore, J. C., & Abou-Khalil, B. (2011). Cross hippocampal influence in mesial temporal lobe epilepsy measured with high temporal resolution functional magnetic resonance imaging. *Epilepsia*, *52*, 1741–1749.

Morris, R. G., Abrahams, S., Baddeley, A. D., & Polkey, C. E. (1995). Doors and People: Visual and verbal memory after unilateral temporal lobectomy. *Neuropsychology*, *9*, 464–469.

Mueller, S. G., Laxer, K. D., Barakos, J., Cheong, I., Garcia, P., & Weiner, M. W. (2009). Subfield atrophy pattern in temporal lobe epilepsy with and without mesial sclerosis detected by high-resolution MRI at 4 Tesla: Preliminary results. *Epilepsia*, *50*, 1474–1483.

Naci, L., Cusack, R., Anello, M., & Owen, A. M. (2014). A common neural code for similar conscious experiences in different individuals. *Proceedings of the National*

Academy of Sciences, 1–6.

Newcombe, F., Ratcliff, G., & Damasio, H. (1987). Dissociable visual and spatial impairments following right posterior cerebral lesions: Clinical, neuropsychological and anatomical evidence. *Neuropsychologia*, *25*, 149–161.

Noachtar, S. (2001). Seizure semiology. In H. O. Luders (Ed.), *Epilepsy: Comprehensive review and case discussions* (pp. 127–140). London: Martin Dunitz.

O'Keefe, J., & Nadel, L. (1978). *The hipocampus as a cognitive map*. London: Oxford University Press.

Palacios Bote, R., Blázquez-Llorca, L., Fernández-Gil, M. Á., Alonso-Nanclares, L., Muñoz, A., & De Felipe, J. (2008). Hippocampal sclerosis: Histopathology substrate and magnetic resonance imaging. *Seminars in Ultrasound, CT and MRI*, *29*, 2–14.

Palombo, D. J., Williams, L. J., Abdi, H., & Levine, B. (2013). The survey of autobiographical memory (SAM): A novel measure of trait mnemonics in everyday life. *Cortex*, *49*, 1526–1540.

Pardoe, H. R., Pell, G. S., Abbott, D. F., & Jackson, G. D. (2009). Hippocampal volume assessment in temporal lobe epilepsy: How good is automated segmentation? *Epilepsia*, *50*, 2586–2592.

Park, L., St-Laurent, M., McAndrews, M. P., & Moscovitch, M. (2011). The immediacy of recollection: The use of the historical present in narratives of autobiographical episodes by patients with unilateral temporal lobe epilepsy. *Neuropsychologia*, *49*, 1171–1176.

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*, 1322–1327.

Paz, R., Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2010). A neural substrate in the human hippocampus for linking successive events. *Proceedings of the National Academy of Sciences*, *107*, 6046–6051.

Pelletier, I., Sauerwein, H. C., Lepore, F., Saint-Amour, D., & Lassonde, M. (2007). Non-invasive alternatives to the Wada test in the presurgical evaluation of language and memory functions in epilepsy patients. *Epileptic Disorders*, *9*, 111–126.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Penfield, W., & Mathieson, G. (1974). Memory: Autopsy findings and comments on the role of hippocampus in experiential recall. *Archives of Neurology*, *31*, 145–154.

Penfield, W., & Milner, B. (1958). Memory deficit produced by bilateral lesions in the hippocampal zone. *Archives of Neurology And Psychiatry*, *79*, 475–497.

Perbal, S., Ehrlé, N., Samson, S., Baulac, M., & Pouthas, V. (2001). Time estimation in patients with right or left medial-temporal lobe resection. *NeuroReport*, *12*, 939–942.

Perbal, S., Pouthas, V., & Van Der Linden, M. (2000). Time estimation and amnesia: A case study. *Neurocase, 6*, 347–356.

Petrides, M. (1985). Deficits on conditional associative-learning tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 23*, 601–614.

Pittau, F., Grova, C., Moeller, F., Dubeau, F., & Gotman, J. (2012). Patterns of altered functional connectivity in mesial temporal lobe epilepsy. *Epilepsia, 53*, 1013–1023.

Placantonakis, D. G., Shariff, S., Lafaille, F., Labar, D., Harden, C., Hosain, S., … Schwartz, T. H. (2010). Bilateral intracranial electrodes for lateralizing intractable epilepsy: Efficacy, risk, and outcome. *Neurosurgery, 66*, 274–283.

Powell, H. W. R., Richardson, M. P., Symms, M. R., Boulby, P. A., Thompson, P. J., Duncan, J. S., & Koepp, M. J. (2007). Reorganization of verbal and nonverbal memory in temporal lobe epilepsy due to unilateral hippocampal sclerosis. *Epilepsia*, *48*, 1512–1525.

Powell, H. W. R., Richardson, M. P., Symms, M. R., Boulby, P. A., Thompson, P. J., Duncan, J. S., & Koepp, M. J. (2008). Preoperative fMRI predicts memory decline following anterior temporal lobe resection. *Journal of Neurology, Neurosurgery and Psychiatry*, *79*, 686–693.

Qin, S., Piekema, C., Petersson, K. M., Han, B., Luo, J., & Fernández, G. (2007). Probing the transformation of discontinuous associations into episodic memory: An event-related fMRI study. *NeuroImage*, *38*, 212–222.

Rabin, M. L., Narayan, V. M., Kimberg, D. Y., Casasanto, D. J., Glosser, G., Tracy, J. I., … Detre, J. A. (2004). Functional MRI predicts post-surgical memory following temporal lobectomy. *Brain*, *127*, 2286–2298.

Ranganath, C., & Hsieh, L. (2016). The hippocampus: A special place for time. *Annals of the New York Academy of Sciences*, *1369*, 93–110.

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia, 42*, 2–13.

Rasmussen, T., & Milner, B. (1977). The role of early left-brain injury in determing lateralization of cerebral speech functions. *Annals of the New York Academy of Sciences*, *299*, 355–369.

Rausch, R., & Babb, T. (1993). Hippocampal neuron loss and mmory scores before and after temporal lobe surgery for epilepsy. *Epilepsy, 50*, 812–827.

Rausch, R., & Walsh, G. O. (1984). Right-hemisphere language dominance in right-handed epileptic patients. *Archives of Neurology, 41*, 1077–1080.

Reminger, S. L., Kaszniak, A. W., Labiner, D. M., Littrell, L. D., David, B. T., Ryan, L., … Kaemingk, K. L. (2004). Bilateral hippocampal volume predicts verbal memory function in temporal lobe epilepsy. *Epilepsy and Behavior, 5*, 687–695.

Richardson, M. P., Strange, B. A., Duncan, J. S., & Dolan, R. J. (2003). Preserved verbal memory function in left medial temporal pathology involves reorganisation of function to right medial temporal lobe. *NeuroImage*, *20*, s112–s119.

Robin, J., & Moscovitch, M. (2017). Details, gist and schema: Hippocampal – neocortical interactions underlying recent and remote episodic and spatial memory. *Current Opinion in Behavioral Sciences*, *17*, 114–123.

Rodin, E., Funke, M., Berg, P., & Matsuo, F. (2004). Magnetoencephalographic spikes not detected by conventional electroencephalography. *Clinical Neurophysiology*, *115*, 2041–2047.

Rosenow, F., & Lüders, H. (2001). Presurgical evaluation of epilepsy. *Brain*, *124*, 1683–1700.

Ross, R. S., & Slotnick, S. D. (2008). The hippocampus is preferentially associated with memory for spatial context. *Journal of Cognitive Neuroscience*, *20*, 432–446.

Sabsevitz, D. S., Swanson, S. J., Morris, G. L., Mueller, W. M., & Seidenberg, M. (2001). Memory outcome after left anterior temporal lobectomy in patients with expected and reversed Wada memory asymmetry scores. *Epilepsia*, *42*, 1408–1415.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage*, *90*, 449–468.

Saling, M. M. (2009). Verbal memory in mesial temporal lobe epilepsy: Beyond material specificity. *Brain*, *132*, 570–582.

Sander, J. W. A. S., Genton, P., & Portera-Sanchez, A. (1993). Some aspects of prognosis in the epilepsies: A review. *Epilepsia*, *34*, 1007–1016.

Sass, K. J., Buchanan, C. P., Kraemer, S., Westerveld, M., Kim, J. H., & Spencer, D. D. (1995). Verbal memory impairment resulting from hippocampal neuron loss among epileptic patients with structural lesions. *Neurology*, *45*, 2154–2158.

Schuele, S. U., & Lüders, H. O. (2008). Intractable epilepsy: Management and therapeutic alternatives. *Lancet Neurology*, *7*, 514–524.

Scoville, W. B. (1954). The limbic lobe in man. *Journal of Neurosurgery*, *11*, 64–66.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*, 11–21.

Seidenberg, M., Hermann, B. P., Schoenfeld, J., Davies, K., Wyler, A., & Dohan, F. C. (1997). Reorganization of verbal memory function in early onset left temporal lobe epilepsy. *Brain and Cognition*, *35*, 132–148.

Semah, F., Picot, M.-C., Adam, C., Broglin, D., Arzimanoglou, A., Bazin, B., … Baulac, M. (1998). Is the underlying cause of epilepsy a major prognostic factor for recurrence? *Neurology*, *51*, 1256–1262.

Shah, P., Bassett, D. S., Wisse, L. E. M., Detre, J. A., Stein, J. M., Yushkevich, P. A., … Davis, K. A. (2019). Structural and functional asymmetry of medial temporal subregions in unilateral temporal lobe epilepsy: A 7T MRI study. *Human Brain Mapping*, *40*, 2390–2398.

Sheldon, S., McAndrews, M. P., & Moscovitch, M. (2011). Episodic memory processes mediated by the medial temporal lobes contribute to open-ended problem solving. *Neuropsychologia*, *49*, 2439–2447.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.

Spencer, S. S. (2003). Neural networks in human epilepsy. *Epilepsia*, *44*, 626–626.

Spiers, H. J., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2001). Bilateral hippocampal pathology impairs topographical and episodic memory but not visual pattern matching. *Hippocampus*, *11*, 715–725.

Spiers, H. J., Burgess, N., Maguire, E. A., Baxendale, S. A., Hartley, T., Thompson, P. J., & O'Keefe, J. (2001). Unilateral temporal lobectomy patients show lateralized topographical and episodic memory deficits in a virtual town. *Brain*, *124*, 2476–2489.

Spreen, O., & Strauss, E. (Eds.). (1991). Rey Visual Design Learning Test (RVDLT). In *A compendium of neuropsychological tests* (pp. 168–176). New York, NY: Oxford University Press.

Spreng, R. N., Mar, R. A., & Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, tavigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, *21*, 489–510.

Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S. F., … Mueller, W. M. (1999). Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Brain*, *122*, 2033–2045.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, *99*, 195–231.

St-Laurent, M., McCormick, C., Cohn, M., Mišić, B., Giannoylis, I., & McAndrews, M. P. (2014). Using multivariate data reduction to predict postsurgery memory decline in patients with mesial temporal lobe epilepsy. *Epilepsy & Behavior*, *31*, 220–227.

St-Laurent, M., Moscovitch, M., Jadd, R., & McAndrews, M. P. (2014). The perceptual richness of complex memory episodes is compromised by medial temporal lobe damage. *Hippocampus*, *24*, 560–576.

St-Laurent, M., Moscovitch, M., Levine, B., & McAndrews, M. P. (2009). Determinants of autobiographical memory in patients with unilateral temporal lobe epilepsy or excisions. *Neuropsychologia*, *47*, 2211–2221.

St-Laurent, M., Moscovitch, M., & McAndrews, M. P. (2016). The retrieval of perceptual memory details depends on right hippocampal integrity and activation. *Cortex, 84*, 15–33.

St-Laurent, M., Moscovitch, M., Tau, M., & McAndrews, M. P. (2011). The temporal unraveling of autobiographical memory narratives in patients with temporal lobe epilepsy or excisions. *Hippocampus, 21*, 409–421.

Staresina, B. P., & Davachi, L. (2009). Mind the gap: Binding experiences across space and time in the human hippocampus. *Neuron, 63*, 267–276.

Stoub, T. R., Chicharro, A. V., Grote, C. L., & Kanner, A. M. (2019). Disconnection of hippocampal networks contributes to memory dysfunction in individuals with temporal lobe epilepsy. *Hippocampus, 29*, 451–457.

Strauss, E., Sherman, E. M. S., & Spreen, O. (Eds.). (2006a). Rey-Osterrieth Auditory Verbal Learning Test (RAVLT). In *A compendium of neuropsychological tests* (pp. 776–810). New York, NY: Oxford University Press.

Strauss, E., Sherman, E. M. S., & Spreen, O. (Eds.). (2006b). Rey Complex Figure Test (ROCF). In *A compendium of neuropsychological tests* (pp. 811–840). New York, NY: Oxford University Press.

Su, L., An, J., Ma, Q., Qiu, S., & Hu, D. (2015). Influence of resting-state network on lateralization of functional connectivity in mesial temporal lobe epilepsy. *American Journal of Neuroradiology, 36*, 1479–1487.

Svoboda, E., McKinnon, M. C., & Levine, B. (2006). The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia, 44*, 2189–2208.

Szaflarski, J. P., Holland, S. K., Jacola, L. M., Lindsell, C., Privitera, M. D., & Szaflarski, M. (2008). Comprehensive presurgical functional MRI language evaluation in adult patients with epilepsy. *Epilepsy and Behavior, 12*, 74–83.

Taylor, J. R., Williams, N., Cusack, R., Auer, T., Shafto, M. A., Dixon, M., … Henson, R. N. (2017). The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *NeuroImage, 144*, 262–269.

Tellez-Zenteno, J. F., Dhar, R., & Wiebe, S. (2005). Long-term seizure outcomes following epilepsy surgery: A systematic review and meta-analysis. *Brain, 128*, 1188–1198.

Tellez-Zenteno, J. F., Pondal-Sordo, M., Matijevic, S., & Wiebe, S. (2004). National and regional prevalence of self-reported epilepsy in Canada. *Epilepsia, 45*, 1623–1629.

Tellez-Zenteno, J. F., Ronquillo, L. H., Moien-Afshari, F., & Wiebe, S. (2010). Surgical outcomes in lesional and non-lesional epilepsy: A systematic review and meta-analysis. *Epilepsy Research, 89*, 310–318.

Thaiss, L., & Petrides, M. (2008). Autobiographical memory of the recent past following

frontal cortex or temporal lobe excisions. *European Journal of Neuroscience*, *28*, 829–840.

Thavabalasingam, S., O'Neil, E. B., Tay, J., Nestor, A., & Lee, A. C. H. (2019). Evidence for the incorporation of temporal duration information in human hippocampal long-term memory sequence representations. *Proceedings of the National Academy of Sciences*, *116*, 6407–6414.

Tracy, J. I., Osipowicz, K., Spechler, P., Sharan, A., Skidmore, C., Doucet, G., & Sperling, M. R. (2014). Functional connectivity evidence of cortico-cortio inhibition in temporal lobe epilepsy. *Human Brain Mapping*, *35*, 1–26.

Tubridy, S., & Davachi, L. (2011). Medial temporal lobe contributions to episodic sequence encoding. *Cerebral Cortex*, *21*, 272–280.

Tulving, E., Markowitsch, H. J., Craik, F. E., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex*, *6*, 71–79.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience*, *30*, 11177–11187.

Vanderwal, T., Eilbott, J., Finn, E. S., Craddock, R. C., Turnbull, A., & Castellanos, F. X. (2017). Individual differences in functional connectivity during naturalistic viewing conditions. *NeuroImage*, *157*, 521–530.

Viskontas, I. V, McAndrews, M. P., & Moscovitch, M. (2000). Remote episodic memory deficits in patients with unilateral temporal lobe epilepsy and excisions. *Journal of Neuroscience*, *20*, 5853–5857.

Voets, N. L., Beckmann, C. F., Cole, D. M., Hong, S., Bernasconi, A., & Bernasconi, N. (2012). Structural substrates for resting network disruption in temporal lobe epilepsy. *Brain*, *135*, 2350–2357.

Voets, N. L., Hodgetts, C. J., Sen, A., Adcock, J. E., & Emir, U. (2017). Hippocampal MRS and subfield volumetry at 7T detects dysfunction not specific to seizure focus. *Scientific Reports*, *7*, 1–14.

Voets, N. L., Zamboni, G., Stokes, M. G., Carpenter, K., Stacey, R., & Adcock, J. E. (2014). Aberrant functional connectivity in dissociable hippocampal networks is associated with deficits in memory. *Journal of Neuroscience*, *34*, 4920–4928.

Voorhies, J. M., & Cohen-Gadol, A. (2013). Techniques for placement of grid and strip electrodes for intracranial epilepsy surgery monitoring: Pearls and pitfalls. *Surgical Neurology International*, *4*, 98.

Wada, J. (1949). A new method for the determination of the side of cerebral speech dominance: A preliminary report on the intracarotid injection of sodium amytal in man. *Igaku Seibutsugaku*, *14*, 221–222.

Wang, F., & Diana, R. A. (2016). Temporal context processing within hippocampal subfields. *NeuroImage*, *134*, 261–269.

Wass, C. T., Rajala, M. M., Hughes, J. M., & Sharbrough, F. W. (1996). Long-term follow-up of patients treated surgically for medically intractable epilepsy: Results in 291 patients treated at Mayo Clinic Rochester between July 1972 and March 1985. *Mayo Clinic Proceedings*, *71*, 1105–1113.

Wechsler, D. (1997). *WMS-III: Wechsler Memory Scale-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). *WAIS-IV: Wechsler Adult Intelligence Scale-Fourth Edition*. San Antonio, TX: Pearson.

Wechsler, D. (2009). *WMS-IV: Wechsler Memory Scale-Fourth Edition*. San Antonio, TX: Pearson.

Wechsler, D. (2011). *WASI-II: Wechsler Abbreviated Scale of Intelligence-Second Edition*. San Antonio, TX: Pearson.

Wellmer, J., von der Groeben, F., Klarmann, U., Weber, C., Elger, C. E., Urbach, H., … von Lehe, M. (2012). Risks and benefits of invasive epilepsy surgery workup with implanted subdural and depth electrodes. *Epilepsia*, *53*, 1322–1332.

Wenger, E., Mårtensson, J., Noack, H., Bodammer, N. C., Kühn, S., Schaefer, S., … Lövdén, M. (2014). Comparing manual and automatic segmentation of hippocampal volumes: Reliability and validity issues in younger and older brains. *Human Brain Mapping*, *35*, 4236–4248.

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*, *2*, 125–141.

Williams, J. M., Medwedeff, C. H., & Haban, G. (1989). Memory disorder and subjective time estimation. *Journal of Clinical and Experimental Neuropsychology*, *11*, 713–723.

Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, *17*, 766–780.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage*, *14*, 1370–1386.

Yang, T., Hakimian, S., & Schwartz, T. H. (2014). Intraoperative electrocorticography (ECoG): Indications, techniques, and utility in epilepsy surgery. *Epileptic Disorders*, *16*, 271–279.

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*, 515–525.

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

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging*, *21*, 466–482.

Zeidman, P., & Maguire, E. A. (2016). Anterior hippocampus: The anatomy of perception, imagination and episodic memory. *Nature Reviews Neuroscience*, *17*, 173–182.

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, *12*, 399–413.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1196–1207.

Appendices

Appendix A: Movie Free and Probed Recall Interview Script

Instruction to examiners: Provide the verbal prompts in quotation marks below.

Free Recall

"I want you to describe the movie clip you just saw. Please provide as much detail as you can about what you saw."

Allow the subject to speak until they have finished without structuring or guiding their response in any way.

General Probe

"Is that all you can tell me about this movie clip? Are there any other details you can remember?"

Specific Probes

"I am going to ask you a few more specific questions about the movie. I do not expect you to remember everything I am going to ask you, because nobody ever does. I will only ask questions to find out whether there is anything else you remember about the movie clip that you haven't thought of telling me. But if you don't remember, just let me know, that's perfectly fine. Please do not guess."

"When did this event take place? Based on what you recall from the movie, could you give an estimate of the…[Year/decade?] [Month/season?] [Time of day?]"

"Where did this take place? Based on what you recall from the movie, could you give an estimate of the… [Country?] [Any specific buildings or settings?]"

“Who were the main characters?” If respondent does not name all 9 speaking characters, add: “Was there anyone else with a speaking part in the movie?”

“What else happened during the course of the movie? Where there any other scenes you now remember that you hadn’t told me about before?”

Perceptual details:

> “Do you remember anything about the décor of the house?”
>
> “What was the mother’s hair colour? What about the father’s?”
>
> “Can you describe the artifact that the uncle brought back?”
>
> “What type of hat was the boy wearing?”
>
> “Were there any specific sounds (not speech) you recall from the movie?”

“Can you tell me anything about what you were thinking or feeling during the film?”

Appendix B: Movie Free and Probed Recall – General Scoring Instructions

Adapting the "Autobiographical Interview Scoring Manual"

In general, the transcript is scored as if the participant is recounting the movie like an autobiographical event.

The "event" as it is referenced in the manual will refer to the movie. Internal details refer to details pertaining to the movie or the participant's experience of the movie. External details are all other details provided.

Details are tallied as per the instructions in the manual; qualitative ratings can be skipped.

Since we are using a structured interview format (for consistency) instead of the intended semi-structured format, it is possible that details provided during probing will have already been given during free recall/general probe. When this happens, do not credit or classify the detail. Only mark these details as repetitions when the information is not specifically probed.

General Scoring Instructions

1. a) Do not give extra credit when the participant elaborates on something that, said more concisely, would earn fewer points. A single memory should only be credited once.

"He was wearing a cowboy hat" = 1 perceptual detail AND "He was wearing a hat. A cowboy hat" = 1 perceptual detail

"It seemed lavish. It wasn't modest." = 1 perceptual detail, 1 repetition

"It wasn't winter. It wasn't fall. I think it was summer = 2 other details, 1 time detail

1. b) When an action is stated in parts, consider whether there is more than one action being reported, or whether the participant is just using additional words/phrases that do not add meaning to the detail.

"He takes off and goes to the supermarket" = 1 event detail

"He left and went home" = 1 event detail

"He walks over to the horse and starts playing on it" = 2 event details (2 separate actions)

2. Do not award less credit to a response that, if said in multiple parts, would earn more credit; do not penalize for being concise.

"He found the gun and the bullets" = 2 event details because, "He found the gun. He also found the bullets" = 2 event details (occurred as separate events in the movie)

<u>BUT</u> "His parents realized he has a gun and bullets" = 1 event detail (occurred as one event in the movie)

3. Credit partial sentences if they convey some meaning (i.e., at least a subject and verb). However, if a participant comes back to this topic, repeats the sentence fragment, and completes it, then score the complete sentence normally and rescore the partial sentence as an "other" detail.

"She…" = no details

"She takes…" = 1 event detail

"She takes…She takes the boy's spot" = 1 other detail and 1 event detail

Scoring Instructions by Detail Category

Event Details

1. The limit of crediting up to 5 characters applies across all recall conditions.

2. Award credit for listing characters separately from the credit awarded for the rest of the sentence/clause (only applicable for the characters' first introduction).

"The girl and her father left the supermarket" = 2 events details (for characters) and 1 place detail OR 1 place detail (if characters have been introduced earlier)

Place Details

1. Do not award credit for an event detail when there is little "action" in the sentence and the main content is the location, despite the example given in the manual (i.e., "We went to the hotel"). Instead, credit this as a place detail. Credit both the event detail and place detail when there is more "action" in the sentence.

"He went to the supermarket" = 1 place detail

"His family sent him out to play around the neighbourhood" = 1 event detail and 1 place detail

2. Positions of *people* made in reference to a stationary *object* should be coded as place details, as these are more likely to represent locations within a room (or analogous positions outdoors). But also use your judgment with respect to whether the object is large enough to constitute a location in a room.

"He stood behind the couch" = 1 place detail

"He knelt behind the tree" or "He knelt next to the horse" = 1 place detail

3. The detail must actually specify a location, however generic, rather than just movement. The location may be implied from context only if the participant has already stated it earlier.

"His uncle came in." OR "He walked out." = 1 event detail

"He exited the room." OR "He went outside." = 1 place detail

4. When the examinee describes a room based on what happened there, code the location and event separately.

"The room where he was unpacking" = 1 place detail, 1 event detail

Time Details

1. References to the beginning or end of the movie are scored as other details.

"In the beginning" OR "In the end" OR "That was the end" = 1 other detail

2. Phrases such as "in the meantime," "at that moment," and "before that" are too vague to reflect specific points in time, and instead convey information about the sequence of events. Therefore, they are coded as event details.

3. In specific probing, only credit years/eras between 1930 and 1969. Years outside this range are considered errors.

4. If you credit "It said Alfred Hitchcock Presents, 1961" (as a perceptual detail), do not also give credit for stating that it was set in the 1960s, as these are not distinct memories. Instead, code the second instance as a repetition.

Perceptual Details

1. Details related to duration (coded as perceptual details) may be awarded credit, even when they are vague.

"Continuously throughout the whole movie…" = 1 perceptual detail

"She was riding for a little while" = 1 event detail and 1 perceptual detail

2. The weather and clothing are listed under event details in the scoring guidelines, but should be coded as perceptual details for the movie.

3. Positions of people or objects made in reference to people or to non-stationary objects should be coded as perceptual details.

“He was next to his mother” = 1 perceptual detail

“His family searched for him in a vehicle” = 1 event detail, 1 perceptual detail

“He found the gun in a suitcase” = 1 event detail, 1 perceptual detail

“The mirror was next to her head” = 1 perceptual detail

4. Visual details related to the bullets in the gun (i.e., their number, position) are coded as perceptual details.

“There was one bullet in the gun” = 1 perceptual detail

“The bullet was next in the chamber” = 1 perceptual detail

5. In response to the artifact probe, credit any new information, even if it is not the desired response (i.e., describing the mask). For example, if the examinee provides descriptive information about the gun that was not previously reported, credit this new information.

Emotion/Thought Details

1. Score emotion and thought details separately, even if they are related to each other.

“I was anxious because I thought he was going to shoot someone” = 2 emotion/thought details (1 emotion and 1 thought)

2. Stating that the clip was suspenseful or caused anxiety should only be credited once during the clip, even if it is felt several times during the clip. Subsequent reports would be counted as repetitions. Variations of these emotions can be credited individually, if they might reasonably reflect different mental states

"It was tense…I was nervous" = 2 emotion/thought details

3. Distinguish between whether the thought is the examinee's opinion (thought detail) or an inference of a character's mental state (event detail).

"The dad was exasperated." = 1 event detail

"The girl was annoying." = 1 thought detail

4. Only credit thought details when you are reasonably certain they had the thought while watching the clip – i.e., it is a memory of a thought they had, not a new thought. The phrase "it seemed" or using past tense is more suggestive of memory, whereas as present tense is more suggestive of a thought after-the-fact. If you cannot be reasonably certain that the thought occurred during the movie, then consider whether one of the external detail categories (e.g., 6 or 8) is a better fit.

"It seemed like something out of the twilight zone" = 1 thought detail

"It was a different time back then I suppose" = 1 semantic detail

"I don't know why he would carry a gun" = 1 other detail

Semantic Details

1. Character names are not awarded extra credit beyond crediting for the character's presence (as an event detail). Instead, character names should be coded as semantic details.

"The uncle was visiting. His name was Rick" <u>OR</u> "The uncle, Rick, was visiting" = 1 event detail, 1 semantic detail

Repetitions

1. Be careful not to discredit reports of the same thing happening in more than one scene. For example, the boy pretends to shoot people several times in different scenes. If it is

clear that the participant is speaking about a different scene, award credit appropriately, rather than coding these as repetitions.

2. a) Synonyms are considered repetitions.

“The maid was there. The housekeeper was there” = 1 event detail, 1 repetition

2. b) Poor or inaccurate synonyms may be credited when the clear intention of the examinee was to provide a synonym.

“[He went on the toy horse,] carousel, [whatever they are called.]” = 1 repetition

2. c) Synonyms are only coded as repetitions if they add no new meaning. If the second term used provides some clarification or elaboration, do not code as a repetition.

“It was a gun. It was a revolver.” = 2 perceptual details

3. Probing is used to determine if any new information can be retrieved from memory. Do not credit details that have already been provided (in any condition), even if they are probed specifically.

Other

1. Phrases like “I think,” “I assume,” “I guess,” “probably,” or “maybe” should not be coded as separate details but rather grouped with the detail that they are referencing. However, use your judgment to consider whether the addition of the clause implies that the detail is an inference and was not specifically recalled. “I think” and “I assume” suggests more confidence than “I guess,” just as “probably” suggests marginally more confidence than “maybe.” Also, details offered during free recall are much less likely to be inferences that details offered during specific probing, and so in general, assume that the detail was not inferred.

[Free recall] “I think maybe he didn’t know it was real.” = 1 event detail

[Free recall] “He loaded another bullet, I guess.” = 1 event detail

[In response to auditory probe] “Maybe the horse made a sound.” = 1 other detail

[In response to boy’s hat probe] “A cowboy hat, I think.” = 1 perceptual detail

2. If the examinee appears to have made an inference, consider whether that information could be directly encoded from the movie, or whether it must be inferred. If it can be directly encoded from the movie and the examinee appeared to be guessing, code as an inference. If the examinee could not directly encode the information from the movie (as in the case of time details, but also country, year, hair colour), then code this information as if it was not an inference.

“The mother’s hair was blonde, I guess” = 1 perceptual detail and 1 other detail

“The boy took the gun, I guess” = 2 other details

Appendix C: Movie Free and Probed Recall – Temporal Scoring Instructions

Only score temporal resolution and order outcome measures based on **correct, internal details.**

Temporal Resolution

General Scoring Instructions

The movie itself is eight-minutes long but it covers an event of several hours, so the eight minutes should be scaled to reflect their temporal resolution on a timescale of several hours (i.e., as if the participant was describing an autobiographical memory of several hours).

Details pertaining to the entirety of the movie (or occasionally multiple scenes) are *temporally indefinite (I)*. E.g.:

"There was a young boy"

"The uncle was visiting."

"It was in black and white"

"The weather was sunny"

Most specific probe responses (e.g., settings, characters, perceptual details) but not necessarily sounds (e.g., "the gun shot" would be clustered)

Scenes can be thought of as "subepisodes," so a detail that spans an entire scene may be considered *temporally precise – higher order (HO)*. It is also possible that a subevent within a scene would fall in this category, as well as a subevent spanning multiple scenes, assuming it has a clear start or end that is distinct from the start/end of the movie. Ask yourself whether the detail is true for the whole

scene (and not the whole movie), and whether the detail can be divided into more specific, smaller actions. If the answer to these questions is 'Yes,' the detail is likely high order. E.g.:

"He went to the grocery store."

"He was riding on the horse."

"His parents were looking for him."

"He thinks it's a toy gun." (starts when he finds it and ends when he shoots it)

A detail that occurs within a single scene, on the order of seconds may be considered *temporally precise – clustered (C).* E.g.:

"He left the room."

"He shot the mirror."

"He put the bullets in his pocket."

Outcome measures: number of temporally indefinite details, number of temporally precise – higher order details, number of temporally precise – clustered details, number of clusters, mean cluster size; sum across all three recall conditions.

Scoring Instructions by Detail Category

Event details

1. "He pointed the gun at people.": Depending on the context, this could represent a clustered or higher order detail. Consider whether the participant is talking about a particular instance of pointing the gun (clustered), or about a scene in general (higher order). For example, "he kept pointing the gun at people" would likely be a higher order

detail; so would reporting it out of context or as part of the specific probing. Reporting it in a sequence with other clustered details would likely make it a clustered detail.

2. Details related to the introduction of characters: coded as "I" (up to the limit of five as per the manual)

Place details

1. Place details provided during free recall and general probe: usually "HO" or "C"

Examples of "HO": "He went back to the house," "He was at the supermarket," "They were next to some trees," "He was on the horse"

Examples of "C": "He left the room," "He entered the room," "He went over to the horse."

2. Place details provided during specific probes: usually "HO" or "I"

Examples of "HO": settings pertaining to one scene (e.g., "bedroom," "backyard")

Examples of "I": "United States," "the suburbs," settings pertaining to multiple scenes (e.g., "living room," "supermarket," "outside")

Time details (usually "I")

1. When the examinee indicates different times of day (e.g., "it starts in the morning…around dinnertime…"): code each time detail as "HO"

2. If the examinee only indicates one time of day during recall, even if it seems to be in reference to one part of the clip (e.g., "it starts in the morning" with no other mention of time of day), assume they are referring to the entirety of the clip and code as "I"

Perceptual details

1. Descriptions of weather, clothing, hair colour, the mask, décor in rooms that span multiple scenes (e.g., living room), noises that are repeated or continuous (e.g., suspenseful music, “pew pew,” rotating the barrel): coded as “I”

2. Descriptions of décor in rooms only shown in one scene (e.g., bedroom), noises that span the length of a scene (e.g., the mechanical horse): coded as “HO”

3. Noises on the order of seconds (e.g., mirror breaking, gun going off, car door shutting, carts moving, etc.): coded as “C”

Thoughts/emotion details (coded as “HO” or “I,” but never “C”)

1. Thoughts/emotions with no temporal context given or associated with recurring events (e.g., pretending to shoot): coded as “I”

2. Thoughts/emotions tied to specific scenes/events that do not recur: coded as “HO”

Temporal Order

Since we know the correct order of details, there is no need to infer the order of events. So a sequencing error is any instance where the details are reported in an order that deviates from their order in the film clip.

Count a sequencing error whether or not the participant recognizes the error. E.g.:

> “The mother was talking to the uncle. Before that, the kids were playing in the yard.” AND “The mother was talking to the uncle, and the kids were playing in the yard” both contain 1 sequencing error.
>
> “Oh, I forgot the part about…” signifies a sequencing error

It is implied that some events occur simultaneously, though the viewer sees them in sequence. In this case it would be correct to order them in one of two ways: (1) as

occurring simultaneously, and (2) as occurring in sequence based on the scene that came first. e.g.:

> **“**The boy was at the supermarket. Meanwhile, the mother was calling around to her friends.” AND “The boy was at the supermarket. Then, the mother called around.” contain no sequencing errors, BUT “The mother called around looking for her son, as he was headed for the supermarket” would be 1 sequencing error.

If there is a break in chronology and the participant produces several details in sequence, only the first detail is considered a sequencing error. Similarly, *if details are provided after the general probe is given, score only the first detail as a sequencing error* (assuming the rest of the details given are provided in order).

Outcome measures: number of details provided out of chronological order; sum across free recall and general probe.

Other examples

“While he was asked to unpack for his uncle, who then left the room with his father, he discovered a service revolver.”

> 1 sequencing error; the participant makes it sound like (a) his uncle left the room after he was asked to unpack for him (correct); and (b) he discovered the revolver while he was asked to unpack (not while he was unpacking; incorrect)
>
> If you give the benefit of the doubt and assume the participant meant that he found the revolver when he was unpacking, then the detail about when they left the room is given out of order so it would still be 1 sequencing error.

Appendix D: Movie Comprehension Questions

No.	Question	Option 1	Option 2
1	What are the two boys pretending to shoot at?	**another boy**	a fence
2	What is the relationship of the visitor to the family?	friend of the family	**boy's uncle**
3	The boy's uncle returned from a trip. What was the reason for the trip?	vacation	**business**
4	What does the boy's father suggest that the boy do for his uncle?	**unpack**	clean
5	When the boy starts unpacking, towards what does he pretend to shoot his toy gun?	the camera	**the mirror**
6	Where does the boy leave his toy gun?	**the suitcase**	the dresser
7	Who was delayed arriving to the house that caused the mother concern?	**the maid**	the boy's friend
8	Who or what did the boy pretend to shoot immediately before he was sent outside to play?	his father	**his mother**
9	What type of mechanical animal does the boy ride?	elephant	**horse**
10	What type of coin does the boy need to activate the mechanical horse?	nickel	**dime**
11	When the boy is on the horse, what does he drop on the ground?	**a bullet**	his hat
12	Where are the adults driving?	**the supermarket**	the police station
13	What does the girl's father give the boy to get him off of the mechanical horse?	a candy bar	**a lollipop**
14	Where does the boy go after leaving the supermarket?	**home**	his friend's house
15	What is the maid doing while she and the boy are talking?	cooking	**setting the table**
16	What is the boy standing behind when he shoots the gun at the maid?	**the couch**	the coffee table
17	What is broken when the boy shoots the gun?	a lamp	**a mirror**

18	What does the boy's uncle pick up after the boy runs to his mother?	the gun	**the mask**
19	In total, how many bullets did the boy put in the gun?	**two**	three
20	When does the boy learn that the gun is real and not a toy?	**after he shoots at the maid**	when he first loads the bullets

Note. Correct answer is shown in bold.

Appendix E: Hippocampal ROI-to-ROI Connectivity in All Participants (n = 67)

Left Hippocampal Connectivity

Targets	beta	*T*	*p*
Right Hippocampus	0.76	25.24	0.000
atlas.Amygdala l	0.48	20.78	0.000
atlas.pPaHC l (Parahippocampal *eft)	0.45	18.47	0.000
atlas.aPaHC l (Parahippocampal *eft)	0.48	18.10	0.000
atlas.Amygdala r	0.41	17.26	0.000
atlas.pPaHC r (Parahippocampal *ght)	0.38	16.07	0.000
atlas.TP l (Temporal Pole Left)	0.33	14.54	0.000
atlas.Cereb45 l (Cerebelum 4 5 Left)	0.32	13.82	0.000
atlas.aMTG l (Middle Temporal G*eft)	0.29	12.09	0.000
atlas.TP r (Temporal Pole Right)	0.27	12.04	0.000
atlas.pTFusC l (Temporal Fusifo*eft)	0.28	11.74	0.000
atlas.MedFC (Frontal Medial Cortex)	0.29	11.62	0.000
atlas.aMTG r (Middle Temporal G*ght)	0.26	11.56	0.000
atlas.pTFusC r (Temporal Fusifo*ght)	0.26	11.06	0.000
atlas.Cereb45 r (Cerebelum 4 5 *ght)	0.25	10.83	0.000
atlas.Thalamus l	0.22	10.71	0.000
atlas.aPaHC r (Parahippocampal *ght)	0.31	10.60	0.000
atlas.aTFusC l (Temporal Fusifo*eft)	0.22	10.13	0.000
atlas.PC (Cingulate Gyrus, post*ion)	0.22	9.99	0.000
atlas.aTFusC r (Temporal Fusifo*ght)	0.21	9.76	0.000
atlas.SubCalC (Subcallosal Cortex)	0.23	9.18	0.000
atlas.pSMG r (Supramarginal Gyr*ght)	-0.19	-8.56	0.000
atlas.PaCiG l (Paracingulate Gy*eft)	0.18	8.46	0.000
atlas.pMTG l (Middle Temporal G*eft)	0.20	8.40	0.000
atlas.FOrb l (Frontal Orbital C*eft)	0.20	8.17	0.000
atlas.TOFusC r (Temporal Occipi*ght)	0.20	8.06	0.000
atlas.PP l (Planum Polare Left)	0.15	7.79	0.000
atlas.aITG l (Inferior Temporal*eft)	0.17	7.53	0.000
atlas.aSMG l (Supramarginal Gyr*eft)	-0.14	-7.39	0.000
atlas.Putamen l	0.15	7.33	0.000
atlas.Ver45 (Vermis 4 5)	0.17	6.84	0.000
atlas.Brain-Stem	0.18	6.64	0.000
atlas.aSMG r (Supramarginal Gyr*ght)	-0.14	-6.52	0.000
atlas.Precuneous (Precuneous Cortex)	0.13	6.50	0.000
atlas.Cereb6 r (Cerebelum 6 Right)	0.15	6.42	0.000
atlas.Cereb9 r (Cerebelum 9 Right)	0.14	6.33	0.000

atlas.PP r (Planum Polare Right)	0.13	6.25	0.000
atlas.Putamen r	0.12	6.13	0.000
atlas.CO r (Central Opercular C*ght)	0.14	6.13	0.000
atlas.Thalamus r	0.14	6.06	0.000
atlas.HG r (Heschl's Gyrus Right)	0.13	6.04	0.000
atlas.aSTG l (Superior Temporal*eft)	0.15	6.01	0.000
atlas.TOFusC l (Temporal Occipi*eft)	0.16	6.00	0.000
atlas.IFG tri l (Inferior Front*eft)	0.12	5.95	0.000
atlas.LG r (Lingual Gyrus Right)	0.13	5.83	0.000
atlas.CO l (Central Opercular C*eft)	0.13	5.77	0.000
atlas.aSTG r (Superior Temporal*ght)	0.13	5.71	0.000
atlas.SCC l (Supracalcarine Cor*eft)	0.10	5.63	0.000
atlas.toITG l (Inferior Tempora*eft)	0.11	5.57	0.000
networks.DefaultMode.MPFC (1,55,-3)	0.33	13.85	0.000
networks.DefaultMode.LP (L) (-3*,33)	0.21	8.94	0.000
networks.Salience.SMG (R) (62,-*,32)	-0.17	-8.14	0.000
networks.Salience.RPFC (R) (32,*,27)	-0.16	-7.49	0.000
networks.DefaultMode.PCC (1,-61,38)	0.14	6.78	0.000
networks.Salience.SMG (L) (-60,*,31)	-0.13	-6.43	0.000
networks.DorsalAttention.IPS (R*,54)	-0.12	-5.88	0.000
networks.DorsalAttention.IPS (L*,52)	-0.11	-5.67	0.000
networks.DefaultMode.LP (R) (47*,29)	0.14	5.57	0.000
atlas.Cuneal r (Cuneal Cortex Right)	0.11	5.51	0.000
atlas.Cereb3 l (Cerebelum 3 Left)	0.11	5.28	0.000
atlas.Ver3 (Vermis 3)	0.13	5.27	0.000
atlas.aITG r (Inferior Temporal*ght)	0.12	5.27	0.000
atlas.LG l (Lingual Gyrus Left)	0.12	5.21	0.000
atlas.pSMG l (Supramarginal Gyr*eft)	-0.11	-5.10	0.000
atlas.pITG l (Inferior Temporal*eft)	0.10	5.06	0.000
networks.Salience.RPFC (L) (-32*,27)	-0.12	-5.14	0.000
atlas.sLOC l (Lateral Occipital*eft)	0.11	5.04	0.000
atlas.Cuneal l (Cuneal Cortex Left)	0.10	5.00	0.000
atlas.IC r (Insular Cortex Right)	0.11	4.96	0.000
atlas.IC l (Insular Cortex Left)	0.13	4.95	0.000
atlas.Ver6 (Vermis 6)	0.11	4.90	0.000
atlas.pMTG r (Middle Temporal G*ght)	0.12	4.79	0.000
networks.FrontoParietal.PPC (R)*,45)	-0.10	-4.70	0.000
atlas.SFG r (Superior Frontal G*ght)	-0.11	-4.64	0.000
atlas.MidFG r (Middle Frontal G*ght)	-0.10	-4.62	0.000
atlas.Cereb3 r (Cerebelum 3 Right)	0.10	4.59	0.000
atlas.HG l (Heschl's Gyrus Left)	0.11	4.56	0.000
atlas.AC (Cingulate Gyrus, ante*ion)	0.11	4.54	0.000

atlas.Cereb9 l (Cerebelum 9 Left)	0.10	4.45	0.000
atlas.ICC l (Intracalcarine Cor*eft)	0.09	4.42	0.000
atlas.PT r (Planum Temporale Right)	0.09	4.41	0.000
atlas.Cereb2 r (Cerebelum Crus2*ght)	0.11	4.36	0.000
atlas.FOrb r (Frontal Orbital C*ght)	0.11	4.31	0.000
atlas.SPL l (Superior Parietal *eft)	-0.08	-4.28	0.000
atlas.Cereb6 l (Cerebelum 6 Left)	0.10	4.28	0.000
atlas.SCC r (Supracalcarine Cor*ght)	0.09	4.19	0.000
atlas.SPL r (Superior Parietal *ght)	-0.09	-4.02	0.000
networks.SensoriMotor.Lateral (*,29)	0.09	4.02	0.000
atlas.Ver9 (Vermis 9)	0.11	4.01	0.000
atlas.Pallidum l	0.08	3.99	0.000
networks.FrontoParietal.LPFC (R*,30)	-0.10	-3.98	0.000
networks.Visual.Medial (2,-79,12)	0.08	3.94	0.000
networks.Language.IFG (L) (-51,26,2)	0.09	3.85	0.000
atlas.IFG oper r (Inferior Fron*ght)	-0.10	-3.77	0.000
networks.SensoriMotor.Lateral (*,29)	0.09	3.71	0.000
atlas.ICC r (Intracalcarine Cor*ght)	0.08	3.68	0.000
networks.Salience.AInsula (R) (*4,0)	-0.07	-3.64	0.001
atlas.Ver12 (Vermis 1 2)	0.08	3.60	0.001
atlas.Accumbens l	0.07	3.42	0.001
atlas.pSTG r (Superior Temporal*ght)	0.08	3.36	0.001
atlas.FP l (Frontal Pole Left)	0.08	3.22	0.002
atlas.PT l (Planum Temporale Left)	0.08	3.19	0.002
atlas.pSTG l (Superior Temporal*eft)	0.07	3.17	0.002
atlas.toMTG l (Middle Temporal *eft)	0.08	3.15	0.002
atlas.FO r (Frontal Operculum C*ght)	-0.07	-3.12	0.003
networks.Visual.Occipital (0,-93,-4)	0.07	3.04	0.003
atlas.Ver8 (Vermis 8)	0.08	3.03	0.004
atlas.Cereb7 l (Cerebelum 7b Left)	-0.06	-2.95	0.004
atlas.pITG r (Inferior Temporal*ght)	0.06	2.90	0.005
atlas.PostCG r (Postcentral Gyr*ght)	0.08	2.78	0.007
atlas.AG r (Angular Gyrus Right)	-0.07	-2.78	0.007
networks.DorsalAttention.FEF (L*,64)	-0.05	-2.73	0.008
atlas.PaCiG r (Paracingulate Gy*ght)	0.07	2.72	0.008
atlas.OFusG r (Occipital Fusifo*ght)	0.06	2.72	0.008
atlas.Accumbens r	0.05	2.70	0.009
networks.SensoriMotor.Superior *,67)	0.07	2.58	0.012
atlas.OFusG l (Occipital Fusifo*eft)	0.06	2.44	0.017
atlas.OP r (Occipital Pole Right)	0.05	2.39	0.020
networks.Language.pSTG (L) (-57*,15)	0.05	2.23	0.029
networks.Cerebellar.Posterior (*-32)	0.06	2.20	0.031

networks.Cerebellar.Anterior (0*-30)	0.06	2.17	0.033

Note. Only significant connections that survive that survive correction for the false discovery rate are shown.

Right Hippocampal Connectivity

Targets	beta	*T*	*p*
atlas.pPaHC r (Parahippocampal *ght)	0.58	26.00	0.000
Left Hippocampus	0.76	25.24	0.000
atlas.Amygdala r	0.53	24.43	0.000
atlas.Amygdala l	0.44	20.63	0.000
atlas.pPaHC l (Parahippocampal *eft)	0.45	18.46	0.000
atlas.pTFusC r (Temporal Fusifo*ght)	0.38	15.91	0.000
atlas.aPaHC r (Parahippocampal *ght)	0.46	15.74	0.000
atlas.Cereb45 l (Cerebelum 4 5 Left)	0.34	14.57	0.000
atlas.aPaHC l (Parahippocampal *eft)	0.35	14.30	0.000
atlas.aTFusC r (Temporal Fusifo*ght)	0.29	13.87	0.000
atlas.Cereb45 r (Cerebelum 4 5 *ght)	0.30	13.81	0.000
atlas.TP r (Temporal Pole Right)	0.31	13.62	0.000
atlas.aMTG r (Middle Temporal G*ght)	0.31	12.43	0.000
atlas.TOFusC r (Temporal Occipi*ght)	0.31	12.42	0.000
atlas.pTFusC l (Temporal Fusifo*eft)	0.27	10.72	0.000
atlas.Thalamus r	0.24	10.68	0.000
atlas.LG r (Lingual Gyrus Right)	0.21	10.46	0.000
atlas.TP l (Temporal Pole Left)	0.25	10.32	0.000
atlas.PC (Cingulate Gyrus, post*ion)	0.24	9.61	0.000
atlas.aTFusC l (Temporal Fusifo*eft)	0.21	9.56	0.000
atlas.Ver45 (Vermis 4 5)	0.22	9.38	0.000
atlas.MedFC (Frontal Medial Cortex)	0.24	8.82	0.000
atlas.aMTG l (Middle Temporal G*eft)	0.21	8.57	0.000
atlas.Putamen r	0.15	8.54	0.000
atlas.Cereb9 l (Cerebelum 9 Left)	0.17	8.21	0.000
atlas.TOFusC l (Temporal Occipi*eft)	0.20	8.06	0.000
atlas.Thalamus l	0.19	8.05	0.000
atlas.Cereb6 r (Cerebelum 6 Right)	0.17	7.98	0.000
atlas.LG l (Lingual Gyrus Left)	0.17	7.88	0.000
atlas.SubCalC (Subcallosal Cortex)	0.21	7.74	0.000
atlas.Cereb3 r (Cerebelum 3 Right)	0.16	7.38	0.000

atlas.aSMG l (Supramarginal Gyr*eft)	-0.15	-7.24	0.000
atlas.Brain-Stem	0.20	7.20	0.000
atlas.IC r (Insular Cortex Right)	0.16	7.19	0.000
atlas.aITG r (Inferior Temporal*ght)	0.18	7.08	0.000
atlas.HG r (Heschl's Gyrus Right)	0.12	7.00	0.000
atlas.FOrb r (Frontal Orbital C*ght)	0.16	6.95	0.000
atlas.PP r (Planum Polare Right)	0.17	6.87	0.000
atlas.Ver3 (Vermis 3)	0.16	6.81	0.000
atlas.Cereb9 r (Cerebelum 9 Right)	0.16	6.67	0.000
atlas.Precuneous (Precuneous Cortex)	0.17	6.65	0.000
atlas.Ver6 (Vermis 6)	0.13	6.61	0.000
atlas.pSMG l (Supramarginal Gyr*eft)	-0.16	-6.56	0.000
atlas.Cereb3 l (Cerebelum 3 Left)	0.13	6.43	0.000
atlas.Cereb6 l (Cerebelum 6 Left)	0.12	6.19	0.000
atlas.aSTG r (Superior Temporal*ght)	0.14	6.12	0.000
atlas.SCC r (Supracalcarine Cor*ght)	0.11	5.87	0.000
atlas.SCC l (Supracalcarine Cor*eft)	0.13	5.78	0.000
atlas.OFusG r (Occipital Fusifo*ght)	0.13	5.75	0.000
atlas.ICC r (Intracalcarine Cor*ght)	0.10	5.59	0.000
atlas.pSMG r (Supramarginal Gyr*ght)	-0.13	-5.57	0.000
networks.DefaultMode.MPFC (1,55,-3)	0.30	12.66	0.000
networks.DefaultMode.LP (R) (47*,29)	0.24	8.64	0.000
networks.Salience.RPFC (L) (-32*,27)	-0.15	-7.30	0.000
networks.Salience.RPFC (R) (32,*,27)	-0.14	-6.99	0.000
networks.DefaultMode.PCC (1,-61,38)	0.17	6.99	0.000
networks.Visual.Medial (2,-79,12)	0.14	6.84	0.000
networks.DefaultMode.LP (L) (-3*,33)	0.18	6.72	0.000
networks.Salience.SMG (L) (-60,*,31)	-0.14	-6.15	0.000
networks.DorsalAttention.IPS (L*,52)	-0.12	-5.92	0.000
networks.Salience.SMG (R) (62,-*,32)	-0.13	-5.86	0.000
atlas.OP r (Occipital Pole Right)	0.13	5.55	0.000
atlas.PP l (Planum Polare Left)	0.11	5.48	0.000
atlas.Cuneal r (Cuneal Cortex Right)	0.11	5.36	0.000
atlas.pMTG r (Middle Temporal G*ght)	0.14	5.32	0.000
atlas.CO r (Central Opercular C*ght)	0.12	5.31	0.000
atlas.PT r (Planum Temporale Right)	0.11	5.28	0.000
networks.Visual.Lateral (R) (38*,13)	0.13	5.51	0.000
atlas.Putamen l	0.10	5.24	0.000
atlas.sLOC r (Lateral Occipital*ght)	0.13	5.21	0.000
networks.Visual.Occipital (0,-93,-4)	0.12	5.17	0.000
atlas.aSMG r (Supramarginal Gyr*ght)	-0.11	-5.03	0.000
atlas.OFusG l (Occipital Fusifo*eft)	0.11	4.94	0.000

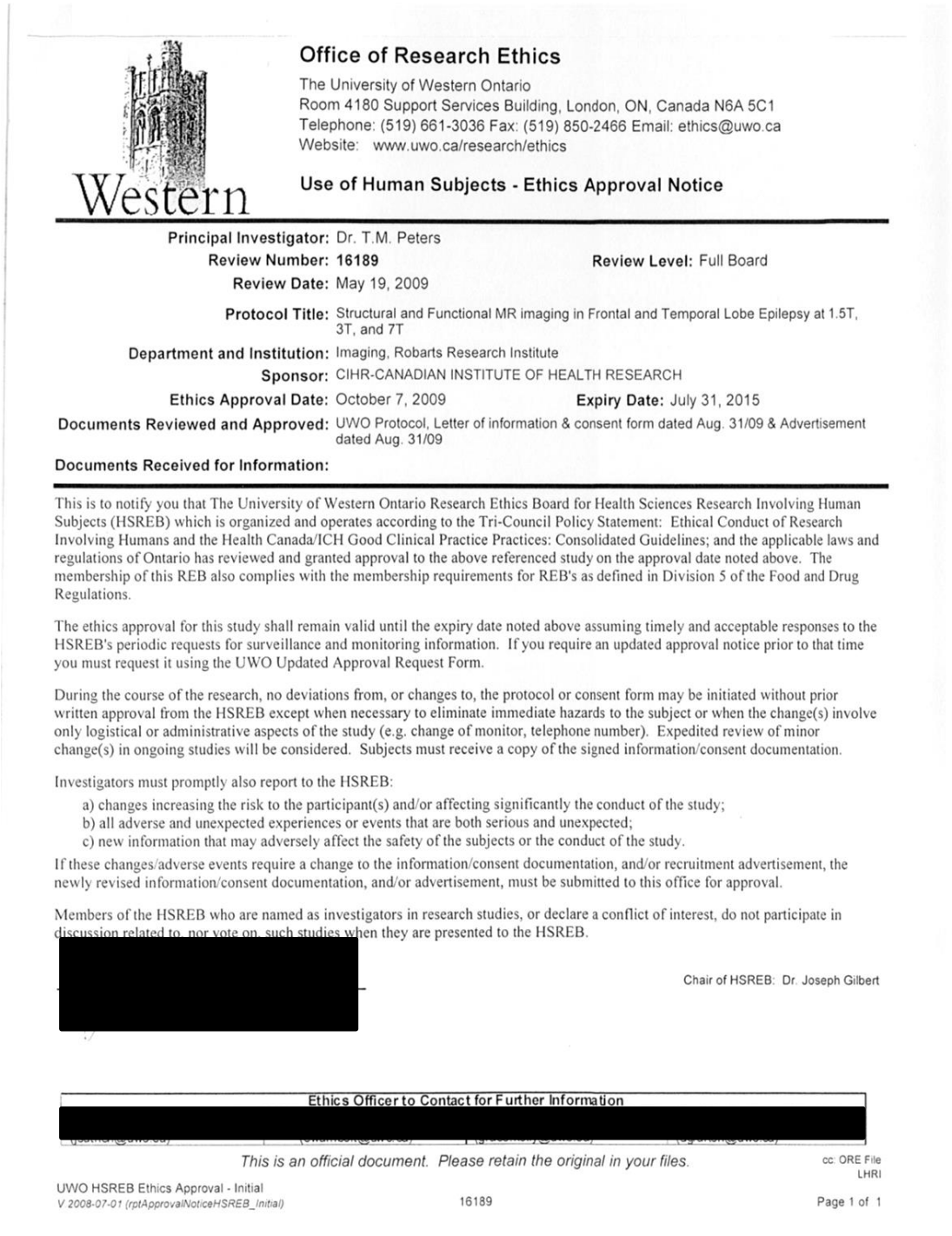
atlas.aITG l (Inferior Temporal*eft)	0.10	4.90	0.000
atlas.ICC l (Intracalcarine Cor*eft)	0.10	4.89	0.000
atlas.Ver12 (Vermis 1 2)	0.10	4.71	0.000
atlas.Cuneal l (Cuneal Cortex Left)	0.10	4.67	0.000
atlas.pITG r (Inferior Temporal*ght)	0.12	4.66	0.000
atlas.toITG r (Inferior Tempora*ght)	0.11	4.65	0.000
atlas.pMTG l (Middle Temporal G*eft)	0.11	4.64	0.000
atlas.Ver9 (Vermis 9)	0.13	4.60	0.000
atlas.iLOC r (Lateral Occipital*ght)	0.11	4.60	0.000
atlas.PaCiG l (Paracingulate Gy*eft)	0.09	4.28	0.000
networks.DorsalAttention.IPS (R*,54)	-0.09	-4.22	0.000
atlas.aSTG l (Superior Temporal*eft)	0.09	4.22	0.000
atlas.CO l (Central Opercular C*eft)	0.09	4.18	0.000
atlas.SPL l (Superior Parietal *eft)	-0.08	-4.04	0.000
atlas.toMTG r (Middle Temporal *ght)	0.11	3.99	0.000
atlas.MidFG l (Middle Frontal G*eft)	-0.08	-3.95	0.000
atlas.OP l (Occipital Pole Left)	0.08	3.83	0.000
atlas.IFG oper l (Inferior Fron*eft)	-0.09	-3.82	0.000
atlas.toITG l (Inferior Tempora*eft)	0.08	3.80	0.000
atlas.sLOC l (Lateral Occipital*eft)	0.09	3.78	0.000
atlas.SPL r (Superior Parietal *ght)	-0.08	-3.54	0.001
atlas.FOrb l (Frontal Orbital C*eft)	0.07	3.50	0.001
atlas.AC (Cingulate Gyrus, ante*ion)	0.08	3.38	0.001
networks.SensoriMotor.Lateral (*,29)	0.06	3.35	0.001
atlas.Ver8 (Vermis 8)	0.08	3.34	0.001
networks.Salience.AInsula (L) (*3,1)	-0.07	-3.28	0.002
atlas.HG l (Heschl's Gyrus Left)	0.07	3.21	0.002
atlas.IC l (Insular Cortex Left)	0.08	3.20	0.002
atlas.PaCiG r (Paracingulate Gy*ght)	0.07	3.04	0.003
networks.Visual.Lateral (L) (-3*,10)	0.07	3.03	0.003
atlas.IFG tri r (Inferior Front*ght)	0.07	3.02	0.004
atlas.FO l (Frontal Operculum C*eft)	-0.06	-3.02	0.004
atlas.SFG r (Superior Frontal G*ght)	-0.06	-3.01	0.004
atlas.pSTG r (Superior Temporal*ght)	0.07	2.95	0.004
networks.FrontoParietal.PPC (L)*,49)	-0.06	-2.86	0.006
atlas.Ver7 (Vermis 7)	0.06	2.81	0.007
atlas.SFG l (Superior Frontal G*eft)	-0.06	-2.80	0.007
networks.DorsalAttention.FEF (L*,64)	-0.06	-2.70	0.009
atlas.Cereb8 r (Cerebelum 8 Right)	0.06	2.69	0.009
networks.FrontoParietal.PPC (R)*,45)	-0.06	-2.63	0.011
networks.Cerebellar.Anterior (0*-30)	0.06	2.54	0.013
atlas.Accumbens r	0.05	2.46	0.016

networks.FrontoParietal.LPFC (L*,28)	-0.06	-2.46	0.017
atlas.Pallidum r	0.04	2.40	0.019
networks.SensoriMotor.Lateral (*,29)	0.05	2.35	0.022
atlas.PT l (Planum Temporale Left)	0.05	2.31	0.024
atlas.FP r (Frontal Pole Right)	0.05	2.17	0.033

Note. Only significant connections that survive that survive correction for the false discovery rate are shown.

Appendix F: Research Ethics Board Approval

Original Ethics Approval for REB #16189 (HSREB #6259)



Western

Office of Research Ethics

The University of Western Ontario
Room 4180 Support Services Building, London, ON, Canada N6A 5C1
Telephone: (519) 661-3036 Fax: (519) 850-2466 Email: ethics@uwo.ca
Website: www.uwo.ca/research/ethics

Use of Human Subjects - Ethics Approval Notice

Principal Investigator: Dr. T.M. Peters
Review Number: 16189 **Review Level:** Full Board
Review Date: May 19, 2009
Protocol Title: Structural and Functional MR imaging in Frontal and Temporal Lobe Epilepsy at 1.5T, 3T, and 7T
Department and Institution: Imaging, Robarts Research Institute
Sponsor: CIHR-CANADIAN INSTITUTE OF HEALTH RESEARCH
Ethics Approval Date: October 7, 2009 **Expiry Date:** July 31, 2015
Documents Reviewed and Approved: UWO Protocol, Letter of information & consent form dated Aug. 31/09 & Advertisement dated Aug. 31/09
Documents Received for Information:

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices: Consolidated Guidelines; and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced study on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the UWO Updated Approval Request Form.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the HSREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of monitor, telephone number). Expedited review of minor change(s) in ongoing studies will be considered. Subjects must receive a copy of the signed information/consent documentation.

Investigators must promptly also report to the HSREB:

a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to this office for approval.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

[redacted]

Chair of HSREB: Dr. Joseph Gilbert

Ethics Officer to Contact for Further Information

[redacted]

This is an official document. Please retain the original in your files.

cc: ORE File
LHRI

UWO HSREB Ethics Approval - Initial
V 2008-07-01 (rptApprovalNoticeHSREB_Initial)

16189

Page 1 of 1

REB Approval for Amendment to Pursue the Present Investigations

Western Research

Research Ethics

Western University Health Science Research Ethics Board
HSREB Amendment Approval Notice

Principal Investigator: Dr. Terry Peters
Department & Institution: Schulich School of Medicine and Dentistry\Medical Biophysics,Robarts Research Institute

Review Type: Full Board
HSREB File Number: 6259
Study Title: Structural and Functional MR imaging in Frontal and Temporal Lobe Epilepsy at 1.5T, 3T, and 7T (REB #16189)
Sponsor: Canadian Institutes of Health Research

HSREB Amendment Approval Date: December 21, 2015
HSREB Expiry Date: May 19, 2016

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Advertisement	Poster for Control Recruitment	2015/11/18
Advertisement	Recruitment Brochure	2015/11/12
Revised Western University Protocol	Updated Protocol Nov 2015 (Clean)	2015/11/13
Revised Letter of Information & Consent	Volunteers	2015/12/14
Revised Letter of Information & Consent	Patients	2015/12/14
Recruitment Items	Poster-Received Dec 9, 2015	2015/12/09
Instruments	Conditional Associative Learning Test	2015/12/18
Instruments	Doors and People - Sample Doors Item	2015/12/18
Instruments	Doors and People - Sample Names Item	2015/12/18
Instruments	Doors and People Administration Instructions	2015/12/18
Instruments	Matrix Reasoning Record Form	2015/12/18
Instruments	Matrix Reasoning Sample Stimulus Pages	2015/12/18
Instruments	Matrix Reasoning, Vocabulary Administration Instructions	2015/12/18
Instruments	RVDLT Administration Chapter (Spreen 1991)	2015/12/18
Instruments	Vocabular Record Form	2015/12/18
Instruments	Vocabulary Sample Stimulus Page	2015/12/18

Western University, Research, Support Services Bldg., Rm. 5150
London, ON, Canada N6G 1G9 t. 519.661.3036 f. 519.850.2466 www.uwo.ca/research/ethics

Western Research

Research Ethics

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the amendment to the above named study, as of the HSREB Initial Approval Date noted above.

HSREB approval for this study remains valid until the HSREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB Continuing Ethics Review.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice Practices (ICH E6 R1), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

[redacted]

Ethics Officer, on behalf of Dr. Joseph Gilbert, HSREB Chair

Ethics Officer to Contact for Further Information: Erika Basile ___ Nicole Kaniki ___ Grace Kelly ___ Mina Mekhail ___ Vikki Tran ___

This is an official document. Please retain the original in your files

Western University, Research, Support Services Bldg., Rm. 5150
London, ON, Canada N6G 1G9 t. 519.661.3036 f. 519.850.2466 www.uwo.ca/research/ethics

Curriculum Vitae

Name:	Daniella (Dora) Ladowski
Post-secondary Education and Degrees:	McGill University Montreal, Quebec, Canada 2007-2011 B.Sc.
	Queen's University Kingston, Ontario, Canada 2012-2014 M.Sc.
	The University of Western Ontario London, Ontario, Canada 2014-pres. Ph.D.
Honours and Awards:	CIHR Master's Award 2012-2013
	Ontario Graduate Scholarship 2013-2014
	NSERC Doctoral Award 2015-2018
Related Work Experience	Predoctoral Neuropsychology Intern, Baycrest Toronto, Ontario, Canada 2018-pres.

Publications:

Ladowski, D., Bowie, C. R., Rodd, J. M., & Johnsrude, I. S. (2018). Use of immediate and more remote context to resolve semantic ambiguities in psychosis. Manuscript submitted for publication.

Ladowski, D., Qian, W., Kapadia, A. N., Macdonald, R. L., & Schweizer, T. A. (2014). Effect of aneurysmal subarachnoid hemorrhage on word generation. Behavioral Neurology, 2014, 1-9.

Banks, S. J., Jones-Gotman, M., Ladowski, D., & Sziklas, V. (2012). Sex differences in the medial temporal lobe during encoding and recognition of pseudowords and abstract designs. NeuroImage, 59, 1888-1895.