Impaired Behavioral Pattern Separation in Refractory Temporal Lobe Epilepsy and Mild Cognitive Impairment.

Permalink
https://escholarship.org/uc/item/32t9s2qz

Journal
Journal of the International Neuropsychological Society : JINS, 28(6)

ISSN
1355-6177

Authors
Lalani, Sanam J
Reyes, Anny
Kaestner, Erik
et al.

Publication Date
2022-07-01

DOI
10.1017/s1355617721000734

Peer reviewed
Impaired behavioral pattern separation in refractory temporal lobe epilepsy and mild cognitive impairment

Sanam J. Lalani\textsuperscript{1}, Anny Reyes\textsuperscript{2,3,4}, Erik Kaestner\textsuperscript{3,4}, Shauna M. Stark\textsuperscript{5}, Craig E.L. Stark\textsuperscript{5}, David Lee\textsuperscript{6}, Leena Kansal\textsuperscript{6}, Jerry J. Shih\textsuperscript{6}, Christine N. Smith\textsuperscript{7,4,8}, Brianna M. Paul\textsuperscript{1}, and Carrie R. McDonald\textsuperscript{4,3,2}

\textsuperscript{1} Department of Neurology, University of California San Francisco, San Francisco, CA, U.S.A.
\textsuperscript{2} San Diego State University, University of California San Diego Joint Doctoral Program in Clinical Psychology, San Diego, CA, U.S.A.
\textsuperscript{3} Center for Multimodal Imaging and Genetics, University of California, San Diego, CA, U.S.A.
\textsuperscript{4} Department of Psychiatry, University of California, San Diego, CA, U.S.A.
\textsuperscript{5} Department of Neurobiology & Behavior, University of California Irvine, Irvine, CA, U.S.A.
\textsuperscript{6} Department of Neurosciences, University of California San Diego, San Diego, CA, U.S.A.
\textsuperscript{7} Veterans Affairs San Diego Healthcare System, San Diego, CA, U.S.A.
\textsuperscript{8} Center for the Neurobiology of Learning and Memory, University of California, Irvine, Irvine, CA, U.S.A.

Author Note

Mailing address: 8950 Villa La Jolla Drive, Multimodal Imaging Laboratory (Suite C101), La Jolla, CA 92037
Telephone: 858-534-2678
Fax:
Email address: camcdonald@health.ucsd.edu

Word count: Full article – 5513/5000; Abstract – 241/250
**Objective:** Episodic memory impairment and hippocampal pathology are hallmark features of both temporal lobe epilepsy (TLE) and amnestic mild cognitive impairment (aMCI). Pattern separation (PS), which enables the distinction between similar but unique experiences, is thought to contribute to successful encoding and retrieval of episodic memories. Impaired PS has been proposed as a potential mechanism underlying episodic memory impairment in aMCI, but this association is less established in TLE. In this study, we examined behavioral PS in patients with TLE and explored whether profiles of performance in TLE are similar to aMCI. **Method:** Patients with TLE, aMCI, and age-matched, healthy controls (HC) completed a modified recognition task that relies on PS for the discrimination of highly similar lure items, the *Mnemonic Similarity Task (MST)*. Group differences were evaluated and relationships between clinical characteristics, California Verbal Learning Test—Second Edition (CVLT-II) scores, and MST performance were tested in the TLE group. **Results:** Patients with TLE and aMCI demonstrated poorer PS performance relative to the HCs, but performance did not differ between the two patient groups. Neither the side of seizure focus nor having hippocampal sclerosis affected performance in TLE. However, TLE patients with clinically-defined memory impairment showed the poorest performance. **Conclusion:** Memory performance on a task that relies on PS was disrupted to a similar extent in TLE and aMCI. The MST could provide a clinically-useful tool for measuring hippocampus-dependent memory impairments in TLE and other neurological disorders associated with hippocampal damage. **Mesh:** episodic memory, memory impairment, cognition, hippocampus, aging, seizure disorder
Introduction

Temporal lobe epilepsy (TLE) is the most common form of focal epilepsy in adults (Engel et al., 2003) and cognitive impairment is a debilitating comorbidity (Bell et al., 2011; Saling, 2009). The neuropsychological profile observed in TLE most frequently involves episodic memory impairment, often characterized by poor encoding of verbal and visual information, depending in part on the side of the seizure focus and the extent of damage to the hippocampus and surrounding medial temporal lobe structures (Bell et al., 2011; Helmstaedter, 2013; Hermann et al., 1997). Although episodic memory impairment in TLE is classically associated with visible hippocampal sclerosis (HS), it is often observed in patients where HS is not visible on MRI according to radiological review (Mueller et al., 2012), supporting evidence that occult hippocampal dysfunction can contribute to impaired memory in TLE.

Episodic memory deficits are also the hallmark feature of amnestic mild cognitive impairment (aMCI), often a prodromal phase of Alzheimer’s Disease that is characterized by poor encoding and retention of verbal and visual information (Griffith et al., 2006; Petersen, 2011). Similar to TLE, the extent of hippocampal atrophy in patients with aMCI has been linked to the degree of episodic memory impairment (Gainotti et al., 2008; Petersen et al., 2000). Thus, although both groups have similar memory profiles and have hippocampal damage, it is not clear whether the mechanisms responsible for memory impairment are the same in these two patient populations (Lapointe et al., 2016; Sen et al., 2018; Tai et al., 2018).

One of the most challenging functions of the episodic memory system is distinguishing between two or more experiences that contain highly overlapping content. This process, known as mnemonic discrimination, has been proposed to rely on an underlying neural mechanism called pattern separation (PS). PS is considered one of the major functions of the hippocampal
circuit and is often impaired in patients with hippocampal damage (Bakker et al., 2008, 2012; Kirwan & Stark, 2007; J. K. Leutgeb et al., 2007; S. Leutgeb et al., 2004; Motley & Kirwan, 2012; Nash et al., 2021; Wais et al., 2017; Yassa et al., 2010). Mechanistically, PS orthogonalizes similar input during encoding, allowing similar memories to be distinguished and co-exist with minimal interference. If the information about similar experiences is successfully orthogonalized, the content can be stored as distinct representations, preserving the uniqueness of each experience. In this case, there is no error or interference from the overlapping content during retrieval of an episode (Tulving, 2002). On the other hand, if similar information is not separated during encoding via PS, memories are less likely to be successfully encoded and accurately retrieved as distinct. In the latter case, recent content overwrites the previously stored similar content, leading to substantial interference and inaccurate memory retrieval (Yassa & Stark, 2011).

Converging evidence from animal and human studies suggests that PS is performed by specific subfields of the hippocampus, specifically the dentate gyrus (DG) and its projections to CA3 (Bakker et al., 2008; J. K. Leutgeb et al., 2007; Wesnes et al., 2014). Because patients with TLE and aMCI are known to have broad hippocampal dysfunction inclusive of the CA3/DG region (Blumcke et al., 2007; Yassa et al., 2010), it is possible that disrupted PS underlies the episodic memory impairments observed in both disorders.

The Mnemonic Similarity task (MST) is a behavioral task that taxes PS (S. M. Stark, Yassa, Lacy, & Stark, 2013). The task involves an incidental encoding phase in which participants view images of common objects followed by a recognition memory task, in which participants distinguish previously viewed “old” objects from “new” objects and from ‘lure’ objects that are novel but “similar” to a previously presented object. In addition to measuring
simple item recognition memory, the MST provides a behavioral metric to evaluate PS because it measures how well participants can distinguish old items from similar lures. An additional manipulation in the task varies the level of difficulty by varying the degree of similarity of the ‘lure’ items (i.e., interference) to the originally presented object. The MST has been widely used to study hippocampus-dependent memory decline associated with normal aging, as well decline due to neurodegenerative processes in patients with aMCI. In healthy aging, recognition memory remains relatively intact with age, while the ability to distinguish old items from similar lures declines with age (Holden et al., 2013; Stark et al., 2013). Furthermore, although patients with aMCI have slightly reduced item recognition memory, they have a disproportionate impairment in tasks that rely heavily on behavioral PS relative to their healthy aging counterparts (Doxey & Kirwan, 2015; Holden et al., 2013; Huffman & Stark, 2017; Stark et al., 2013, 2015, 2019; Stark & Stark, 2017; Toner et al., 2009).

Little is known about behavioral PS in TLE, as no studies have examined behavioral PS in TLE using the MST—the prototypic task of PS in humans. Two published studies have addressed performance that depends on successful PS to date, one using a spatial interference task (Reyes et al., 2018) and one using a semantic interference task (Poch et al., 2020). Both studies demonstrated that PS was impaired in patients with TLE compared to age-matched controls, and that high levels of interference (i.e., highly similar lures) during encoding were associated with worse PS. These initial studies highlighted the importance of impaired performance on tasks that tax PS ability in our understanding of memory dysfunction in TLE. However, studies in TLE utilizing the MST, specifically, are warranted given that the MST is emerging as a clinical tool for testing memory in clinical trials of aMCI (Bakker, Albert, Krauss, Speck, & Gallagher, 2015; Bakker et al., 2012; Papp et al., 2020) and it has recently been
integrated into computerized testing platforms (Cogstate, 2020). Moreover, no studies have directly compared profiles of performance on tasks that measure PS in TLE to those observed within other patient groups with known hippocampal damage (e.g., aMCI). This type of comparison is important as it addresses whether different performances on the MST are universal when hippocampal dysfunction is present or whether they are unique to other aspects of a given disease. To that end, the current study aims to explore PS performance in TLE using the MST to better characterize deficits in this group. We hypothesize that patients with TLE will show impaired behavioral PS relative to age-matched controls and that impairment will be particularly pronounced for TLE patients with HS. Given similarities in the affected brain regions between TLE and aMCI, we further hypothesize that the profile of behavioral PS performance in TLE will be similar to that of aMCI.
Methods

Participants

**Temporal Lobe Epilepsy**

This study was approved by the Institutional Review Boards at UC San Diego and UC San Francisco. Informed consent was collected from all participants in accordance with the Declaration of Helsinki. The current study included 28 patients diagnosed with medically refractory TLE (see Table 1) by a board-certified neurologist with expertise in epileptology, in accordance with the criteria defined by the International League Against Epilepsy (Fisher et al., 2017), and based on evidence from video-EEG, seizure semiology, and neuroimaging evaluation that was consistent with TLE. Patients were undergoing a pre-surgical evaluation at one of the two epilepsy centers (UC San Diego or UC San Francisco) and were recruited if they were between the ages of 18-55 years old. Patients suspected to have TLE were tested for the study (N = 52) and excluded if the seizure origin was ultimately determined to be multi-lobar (n = 9) or extratemporal (n = 1), or if the evaluation was nondiagnostic for a temporal lobe seizure focus (n = 10). Four patients were excluded for other reasons (e.g., nonfocal seizures or prior surgeries). Of the remaining 28 patients, 11 had HS (see Table 2) based on visual inspection of T1 and fluid-attenuated inversion recovery (FLAIR) imaging by board-certified neurologists as part of the pre-surgical evaluation process.

[INSERT TABLE 1 and 2]

**Amnestic Mild Cognitive Impairment**

For comparison with TLE, 18 individuals with aMCI were included in the study. These participants were drawn from an established population at the University of California Irvine (UCI) Alzheimer’s Disease Research Center (ADRC) and all were recruited and diagnosed
according to the same criteria. In brief, a single-domain aMCI diagnosis was based on clinical consensus which required a Clinical Dementia Rating (CDR) score of 0.5 (Hughes et al., 1982) and conformity with the Petersen criteria (Petersen et al., 1999), which include subjective memory complaint corroborated by an informant; objective memory impairment (i.e., when assessed clinically, 1.5 standard deviations or more below the normative group) with otherwise intact cognitive function and intact basic activities of daily living without dementia. Performance on the MST for 11 of the 18 aMCI participants has been reported elsewhere (Stark et al., 2013).

**Healthy Control Groups**

Two healthy control (HC) groups were also drawn from an existing UCI sample. Twenty-four young to middle-aged HCs were age-matched to the TLE participants (Young-HC group), while a sample of 19 older HCs, greater than 60 years of age, served as a control group for the aMCI group (Older-HC; see Table 1). Based on a full neuropsychological evaluation (see Supplementary Table 1), these participants performed within the normal range for their age group (Stark et al., 2013).

**Materials**

**Mnemonic Similarity Task**

All participants completed the MST, an object recognition task (Figure 1) that consists of color photographs of common objects presented in two phases. In the first phase, participants engage in an incidental encoding task as they view an object and make a judgement with a button press to identify an item as an “Indoor” or “Outdoor” item (128 items total, displayed for 2 seconds each, 0.5 second interstimulus interval). The second phase of the task begins immediately thereafter as a surprise recognition memory test and consists of 192 images. The participant was asked to identify each item as “Old,” “Similar,” or “New.” One-third of the items
were identical to the items presented during the study phase (i.e., targets), one-third were novel items that had not been presented during the study phase (i.e., foils), and the remaining items were novel items that were similar but not identical to those seen presented during the encoding phase (i.e., lures).

[INSERT FIGURE 1]

The MST provides two main behavioral performance measures of interest. The first measure is the Lure Discrimination Index (LDI), which is unique to the MST. The LDI depends heavily on PS because it reflects the ability to resolve interference and make the fine-grained distinctions when identifying a similar item from an exact repetition. Specifically, it reflects accuracy for identifying lures as “Similar” rather than mistakenly identifying them as targets by calling them “Old.” The LDI is calculated as the rate of “Similar” responses to lure items (correct trials) minus the rate of “Similar” responses to foils (incorrect trials). In this way, the LDI provides a proportion correct measure that accounts for any bias in designating items as “Similar.” The second measure is a measure of traditional recognition memory discriminability defined as the rate of “Old” responses to target items (i.e., hits) minus the rate of “Old” responses to foils (i.e., false alarms). In this way, the recognition memory measure provides a proportion correct measure that accounts for any bias in designating items as “Old.”

In addition to the LDI, the demand on PS ability is parametrically manipulated in the MST by presenting lure items with varying levels of similarity (Lacy et al., 2011), that is from one of five lure bins, each with a different degree of lure similarity. This manipulation effectively allows for an examination of memory as a function in increasing reliance on PS. As item overlap increases, that is, as lure items become more similar to targets, successful performance requires increasing PS capacity so that encoded items can be distinguished from
similar lures during retrieval (see Figure 1B). Items from the first lure bin (Lure 1 in Figure 1) are the most similar and therefore the most difficult to mnemonically distinguish from target items. Lure items from the last bin (Lure 5) are the least similar and are the easiest to mnemonically distinguish from target items. Lure items were evenly distributed amongst the five lure bins with approximately 13 trials per bin. Response accuracy as a function of lure similarity is explored by examining success in correctly designating lures as “Similar”; calculated as 1 minus the probability of labelling a lure as “Old.”

**Clinical Episodic Memory Measure**

Out of the 28 TLE patients, 16 patients were given the California Verbal Learning Test Second Edition (CVLT-II) to assess episodic memory performance (Delis et al., 2000). Impairment was defined as 1.5 standard deviations or more below the mean of the normative sample on the Long Delay Free Recall trial (LDFR). Seven patients were labeled as impaired on LDFR (TLE-I) and the remaining 9 patients as not impaired (TLE-NI). Of the seven TLE-I, four were HS+ and of the nine TLE-NI, three were HS+.

**Statistical Analyses**

T-tests were conducted to test for differences in age and Fisher’s exact tests were used to examine differences in sex distribution between each patient group (TLE or aMCI) and their age-matched HC group. Analyses of variance (ANOVA) was used to test for differences in LDI and recognition discriminability scores across groups (TLE, aMCI, Young-HC, and Older-HC). When results from the ANOVA were significant, group differences were assessed using post-hoc pairwise tests with Bonferroni correction. Additionally, to analyze performance in patients with TLE and aMCI relative to their age-matched HCs across varying levels of lure similarity, two 2 x 5 mixed ANOVA models were used with group (TLE versus Young-HC or aMCI versus Older-
HC) as the between-subjects factor and lure bin (L1, L2, L3, L4, L5) as the within-subjects factor. This initial ANOVA was performed across all four groups to enable full examination of the effects of both age and disease (TLE vs aMCI) on MST performance. The secondary ANOVA was performed to remove the influence of age allowing us to better characterize the influence of disease alone on performance patterns. To accomplish the latter goal, individual z-scores were calculated for the patient groups based on the mean and SD from their corresponding age-matched HC group. Z-scores were then averaged to obtain the group means for the TLE and aMCI groups (Figure 3; panel C). Nonparametric tests (Mann-Whitney U, Kruskal-Wallis) were used to test for differences in LDI and recognition discriminability within clinical subgroups of TLE participants (HS+ and HS-; TLE-I and TLE-NI; and by side of seizure onset: L TLE, R TLE, B TLE). When relevant, Spearman’s Rho correlations were conducted to evaluate the relationship between MST metrics (LDI and recognition discriminability) and the TLE clinical variables.
Results

Participant Demographic

There was no difference in age between TLE and Young-HC ($t[50] = .795, p = .431$) or between aMCI and Older-HC ($t[35] = 1.29 p = .204$). The groups did not differ in sex composition ($FE = 2.31, p = .523$). Within the TLE group, neither age of onset nor disease duration was significantly correlated with the CVLT, LDI, or recognition discriminability.

Group Differences in MST Performance

There were differences in the LDI across groups ($F[3, 85] = 12.29, p < .001, \eta^2 = 0.30$), as illustrated in Figure 2A. Post hoc comparisons revealed patients with TLE performed worse than Young-HC ($p < .001; \text{Cohen’s } d = -1.36$). The aMCI group performed worse than both the Young-HC group ($p < .001; \text{Cohen’s } d = -1.74$) and the Older-HC group ($p = .045; \text{Cohen’s } d = -.962$). Consistent with our hypothesis, TLE patients performed similarly to the aMCI group ($p = 1.00; \text{Cohen’s } d = .233$) and the Older-HC group ($p = .165; \text{Cohen’s } d = -.649$). As expected, the prior finding that older adults perform worse than young adults on LDI (Stark et al., 2013) was obtained for the Older-HC and Young-HC groups ($p = .017; \text{Cohen’s } d = .767$), given that these participants were obtained from that study. Accordingly, differences between the HC groups will not be examined further.

There were also differences in recognition discriminability across groups ($F[3, 85] = 13.35, p < .001, \eta^2 = 0.32$) as illustrated in Figure 2B. Post hoc comparisons revealed patients with TLE performed worse than both young-HC ($p < .001; \text{Cohen’s } d = -1.29$) and older-HC ($p < .001; \text{Cohen’s } d = -1.70$). The aMCI group also performed worse than the Young-HC group ($p = .027; \text{Cohen’s } d = -1.25$) and the Older-HC group ($p = .003; \text{Cohen’s } d = -1.88$). Consistent with
our hypothesis, recognition discriminability did not differ between the TLE and aMCI groups (p = .933; Cohen’s $d = -0.35$).

[MISC FIGURE 2]

**MST Performance By Level of Similarity**

**TLE and Young-HC**

The 2 (group) x 5 (lure similarity) mixed ANOVA with a Greenhouse-Geisser correction revealed a significant main effect of lure similarity on performance ($F [3.419, 170.926] = 55.576, p < .001$; see Figure 3A); however, the group by lure similarity interaction was also significant ($F [3.419, 170.926] = 3.88, p = .007$), suggesting that the effect of lure similarity on performance differed between the groups. Planned polynomial contrasts revealed a significant linear effect of lure similarity on performance ($F [1, 50] = 145.445, p < .001$), and this linear trend varied by group ($F [1, 50] = 9.96, p = .003$). These findings indicate that performance improved as lure similarity decreased.

**aMCI and Older-HC**

The 2 (group) x 5 (lure similarity) mixed ANOVA with a Greenhouse-Geisser correction revealed a significant main effect of lure similarity on performance, ($F [3.579, 121.683] = 31.624, p < .001$; see Figure 3B); however, the interaction of lure similarity by group was significant ($F [3.579, 121.683] = 2.837, p = .027$), suggesting that the effect of lure similarity on performance differed between the groups. Planned polynomial contrasts revealed a significant linear effect of lure similarity on performance ($F [1, 34] = 120.574, p < .001$), and this linear trend varied by group ($F [1, 34] = 10.123, p = .003$). Similar to TLEs, these findings indicate that performance improved as lure similarity decreased.

**TLE and aMCI**
Group comparisons of the z-scores for TLE and aMCI participants at each level of lure similarity revealed no significant differences between the groups across lure bins 1 through 5 (all p-values > .05). However, there were nuanced differences between the TLE and the aMCI groups in their response style across the three trial types (Table 3; i.e., target, lure, and foil). The aMCI were more likely to identify items as “Old” (t[44] = -2.26, p = .03, Cohen’s $d = 0.11$), while TLE were more likely to identify items as “Similar” (t[44] = 3.56, p = .001, Cohen’s $d = 0.12$). The propensity to identify items as “New” was not significantly different between the patient groups (t[44] = -1.98, p > .05).

[INSERT FIGURE 3]

**MST performance by TLE subtypes**

Neither LDI nor recognition discriminability differed between TLE patients who were HS+ versus HS- (Figure 4A; LDI: $U = 77.5$, $p = .46$, Cohen’s $d = 0.14$; recognition discriminability: $U = 80.0$, $p = .55$, Cohen’s $d = 0.10$) or among patients with right, left or bilateral seizure onset (Figure 4B; LDI: $H(2) = .83$, $p = .66$, $\eta^2 = -0.05$; recognition discriminability: $H(2) = .94$, $p = .62$, $\eta^2 = -0.04$). The TLE-I group performed worse on the LDI relative to the TLE-NI group (Figure 4C; $U = 4$, $p = .002$, Cohen’s $d = 1.83$), even though these two groups performed similarly on recognition discriminability ($U = 25.5$, $p = .536$, Cohen’s $d = 0.40$). Indeed, higher CVLT-II LDFR performance (TLE-NI) was associated with better LDI scores ($r_s = .508$, $p = .04$); there was no relationship with recognition discriminability ($r_s = .041$, $p = .88$).

[INSERT FIGURE 4]

Because the LDI was worse in TLE-I compared to TLE-NI patients at the group level, we tested the sensitivity of the LDI to episodic memory impairment at the single-subject level. A
post-hoc stepwise linear classifier that included LDI and recognition discriminability
demonstrated that the LDI alone correctly discriminated TLE-I patients from TLE-NI ($\chi^2 [1, N = 16] = 8.43, p = .004$) with 75% accuracy (75% cross-validation), 86% sensitivity, and 67% specificity.
Discussion

We examine performance on a task designed to measure behavioral PS in patients with TLE and compared their performance to patients with aMCI, another patient group with known hippocampal damage. We also examined how epilepsy-related clinical characteristics affected LDI performance. We demonstrated a significant reduction in LDI performance in individuals with TLE was similar in magnitude to that observed in older adults with aMCI. Though LDI performance did not differ between patients with right or left TLE or between those with or without HS, patients with TLE who showed a clinically-defined episodic memory impairment demonstrated poorer LDI scores than those with intact episodic memory. These findings suggest that disrupted PS may in part, underlie episodic memory impairments in TLE, and that the MST may be a clinically useful tool for detecting impaired memory encoding related to hippocampal dysfunction at the group and individual subject levels.

Pattern separation impairment in TLE

Despite an increasing number of studies demonstrating impaired PS performance in normal aging and aMCI, few studies have examined PS performance in patients with TLE and no studies have examined PS in TLE using the MST. In previous work from our lab, we used a delayed match-to-sample task with varying levels of spatial interference and showed reduced spatial PS performance in patients with TLE compared to age-matched controls (Reyes et al., 2018). We also demonstrated that impairment under high, but not low, spatial interference was associated with hippocampal volume loss and visuospatial memory impairment. In a second study, Poch and colleagues (2020) used a visual (object) mnemonic discrimination task that varied in the similarity and number of exemplars per object category. They also found reduced PS performance in patients with TLE compared to controls but noted that the differences were
observable only when more exemplars from a category were stored in memory (i.e., high memory interference). Collectively, these studies provide additional support that disrupted PS performance may contribute to the episodic memory impairment observed in TLE and that this process is invariant to the types of stimulus type (i.e., spatial vs. object) and task design.

**The relationship between PS, TLE, and aMCI**

Unlike prior studies, we quantified the magnitude of PS impairment in TLE by comparing performance to several comparison groups. As expected, both TLE and aMCI groups showed reduced LDI and recognition discriminability scores compared to their age-matched healthy counterparts. Importantly, TLEs and aMCIs showed a similar degree of impairment, even as memory interference increased. These findings suggest that these broad profiles of performance are not age or disease-specific but are common to groups with hippocampal dysfunction.

However, the TLE and aMCI groups showed nuanced differences in their response patterns, suggesting disease-specific features that could be clinically meaningful. Regardless of trial type, patients with TLE were more likely to endorse items as “Similar,” whereas patients with aMCI were more likely to endorse items as “Old.” While the latter response bias has previously been reported in healthy aging (Stark et al., 2013) in a sample of younger amnestic patients similar in age to our TLE sample, there was also a greater propensity to misidentify items as “Similar” (Kirwan et al., 2012). This suggests that the difference in response bias observed in TLE versus aMCI may be only partially influenced by aging. Another hypothesis that may explain this response bias is that patients with TLE may rely more on familiarity judgements to make decisions, whereas patients with aMCI are more prone to false positive responses. Data supporting this hypothesis comes from studies revealing that patients with TLE
who experience a common seizure semiology (i.e., ictal déjà vu) tend to make inaccurate familiarity judgements during recognition testing, whereas false positive errors are more common in patients with aMCI (Schacter et al., 1997, 1998), including mistakenly identifying lure items as having been previously encountered (Doxey & Kirwan, 2015; Koutstaal et al., 1999; Koutstaal & Schacter, 1997; Toner et al., 2009).

Neuroanatomically, although there is evidence that CA3/DG can be affected in both aMCI and TLE (Hatanpaa et al., 2014; Mueller et al., 2010; Mueller et al., 2012; Yassa et al., 2011; Yassa et al., 2010), CA1 is often the most vulnerable region in aMCI (de Flores, et al., 2015), whereas cell loss in TLE patients typically includes CA1 but is also quite variable, depending on the HS subtype (Blumcke et al., 2007). These neuropathological differences between the patient groups could explain differences in their response patterns. In addition, the multiple subtypes of HS within TLE that could lead to variability in performances even within TLE patients (Blumcke et al., 2007; Coras et al., 2014). Finally, recent data suggest the recruitment of widespread extra-temporal networks during tasks of PS that may be differentially affected in aMCI and TLE which could lead to slight differences in MST performance (Nash et al., 2021). Thus, poor LDI performance in these two patient populations appears superficially similar but may differ due to underlying, disease-specific pathology.

The relationship between TLE clinical characteristics and PS performance

The present study was also the first to explore how clinical characteristics of the TLE sample influenced behavioral PS performance when measured by the MST. Contrary to our hypothesis, patients with or without HS performed similarly. There are several possible explanations for this finding. First, the sample size for these groups was small and we may have been underpowered to detect this difference. The effect size of the difference between these
groups was small, indicating that a significant difference might be identified if the groups were larger. Second, the MST may be able to detect subtle memory impairments that do not appear as a result of HS on radiological review. For example, both HS+ and HS- groups were impaired relative to controls, suggesting that hippocampal dysfunction due to TLE was sufficient to impair performance on the MST, regardless of whether visible HS was also present. It is worth noting that approximately 40% of TLE patients with normal-appearing hippocampi based on visual inspection of MRIs have been found to have evidence of HS on histopathological analysis (Capraz et al., 2015). Therefore, many patients labeled as HS- may have mild HS that was not detected by visual inspection of MRIs.

When separating patients by impairment on the CVLT-II, we found the CVLT-impaired patients were also impaired on LDI but were comparable to unimpaired patients on the measure of recognition discrimination memory. In addition, performance on the MST was not affected by side of seizure onset. The right, left, and bilateral TLE groups all showed impaired LDI performance relative to controls, but these groups did not differ from one another. This may suggest that PS performance depends upon the integrity of both hippocampi. Although no studies have tested this directly, bilateral hippocampal involvement has been inferred from animal (Gilbert et al., 1998; J. K. Leutgeb et al., 2007) and human (Duff et al., 2012; Kirwan et al., 2012) studies involving damage to both hippocampi, as well as from fMRI studies revealing recruitment of both hippocampi during PS task performance (Motley & Kirwan, 2012). Bilateral involvement may also reflect the fact that the MST includes dually-encodable stimuli (i.e., visual objects that can be verbalized) with lures that are similar in terms of both conceptual and visuoperceptual properties. However, there are also data to suggest that the specific task demands or the strategies employed may lead to greater recruitment of one hippocampus over the
other. For example, using a semantic strategy may place greater demands on the left hippocampus (Doxey & Kirwan, 2015), whereas employing a spatial strategy may place greater demands on the right hippocampus (Duff et al., 2012; Motley & Kirwan, 2012). Together, these data provide evidence that impaired PS may be a mechanism underlying episodic verbal memory impairments in TLE that is sensitive to subtle hippocampal dysfunction (e.g., in HS-patients) and may rely on bilateral hippocampal functioning.

The MST has generated interest as a possible clinical tool for measuring memory in several clinical populations including multiple sclerosis (Planche et al., 2017; Zuppichini & Sandry, 2018), brain tumor (Shiroma et al., 2016), schizophrenia (Martinelli & Shergill, 2015), first episode psychosis (Kraguljac et al., 2018) and autism (South et al., 2015). Advantages of the MST are that it is relatively brief and easy to administer, adaptable (i.e., difficulty of lure bins can be adjusted in order to manipulate the level of interference during encoding), it has multiple stimulus sets for repeat testing, and it provides both a behavioral PS metric as well as a more traditional recognition discriminability metric. Thus, the MST provides clinicians with a flexible memory test in comparison to traditional neuropsychological tests of episodic memory that often have ceiling or floor effects. Variations of the MST have already been incorporated into commonly-used computerized batteries, (i.e., Cogstate) and are used clinically (Cogstate, 2020). The MST is now being used in clinical trials with aMCI and Alzheimer’s Disease to study memory function, assist with diagnosis, and measure treatment response (Bakker et al., 2012, 2015; Papp et al., 2020). Although it is not clear whether the MST will have greater clinical utility than traditional neuropsychological measures of episodic memory, our results provide initial data for the ability of the MST to detect memory impairment in TLE at both the individual and group level. Coupled with the advantages listed above, this measure may also provide a
useful tool for testing the effects of medical or surgical interventions on memory and hippocampal functioning in epilepsy. However, its clinical utility in epilepsy must be empirically established in future studies.

**Limitations**

There are several limitations to this study that should be addressed. First, the number of patients included in both patient samples is modest and limited our ability to run many sub-group comparisons. Second, our sample included only patients with drug-resistant TLE. Therefore, it is not clear if our results would generalize to more benign TLE samples; however, TLE patients were recruited from two different locations which allows for additional generalizability of the results. Third, consistent neuropsychological data were only available for a portion of the patients due to the fact that evaluations were performed in clinical settings using a variety of neuropsychological batteries. Access to raw test data was also limited. As a result, we were not able to explore the relationship between performance on the MST and more specific memory indices on the CVLT-II or other neuropsychological domains. In particular, examining the association between MST performance and List B intrusion errors on the CVLT would be of interest since both increase interference during encoding. Furthermore, even though the main findings for the MST reflect deficits in memory, recent evidence suggests that MST performance may be sensitive to non-mnemonic processes (Davidson et al., 2019; Foster & Giovanello, 2020; Nash et al., 2021; Pishdadian et al., 2020). Exploring the relative contribution of mnemonic and non-mnemonic cognitive abilities (e.g., visuospatial processes, processing speed, inhibitory control) on MST performance will help to determine the relative contribution of these non-memory abilities. Given the verbal and visual nature of the task stimuli, examining the association between MST performance and non-verbal memory is also warranted. Fourth,
pattern completion is a complementary process to PS that allows stored memories to be retrieved with partial or degraded cues (Rolls, 2013). Although we did not include a parallel test of pattern completion in our study, pattern completion deficits may have also contributed to poor MST performance and/or episodic memory impairments (e.g., Baker et al., 2016). Therefore, including both types of measures in future studies will allow better characterization of the source of episodic memory impairments in TLE. Finally, we did not have access to MRI scans which precluded our ability to relate our behavioral findings back to quantitative MRI volumes and features. Future studies, particularly those with high-resolution T2-weighted structural MRI scans, will be able to quantify the volumes of individual hippocampal subfields. These studies will be better positioned to examine the associations between MST performances and hippocampal subfield volumes for a richer neurobiological explanation of the present results.

Conclusions

The current study is the first to utilize the MST in a multi-center setting to identify behavioral PS impairments in TLE, explore the relationship between performance and patient clinical characteristics, and compare performance profiles with another population with hippocampal dysfunction (i.e., older adults with aMCI). We demonstrate similar impairment profiles in both patient groups, further supporting the involvement of hippocampal dysfunction, rather than disease-specific factors, in impaired PS performance in TLE. Future work should be conducted to establish whether the MST offers a more sensitive or specific measure of memory dysfunction compared to existing measures (e.g., CVLT-II), including whether it can detect subtle memory impairment in TLE, detect pathology within specific hippocampal subfields, and/or estimate risk for postoperative memory decline.

Acknowledgements
This work was supported by the National Institute of Health / National Institute of Neurological Disorders and Stroke (C.R.M., R01 NS065838, E.K., T32 MH018399, A.R., F31 NS111883-01). There are no conflicts of interest to report.
Table 1

*Participant Demographics*

<table>
<thead>
<tr>
<th></th>
<th>TLE</th>
<th>Young-HC</th>
<th>aMCI</th>
<th>Older-HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>28</td>
<td>24</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>Age, years, mean (SD)</td>
<td>33.5 (8.4)</td>
<td>31.7 (8.1)</td>
<td>75.3 (9.7)</td>
<td>71.6 (7.3)</td>
</tr>
<tr>
<td>Age, years, range</td>
<td>20-49</td>
<td>20-48</td>
<td>55-93</td>
<td>60-87</td>
</tr>
<tr>
<td>Sex, F/M</td>
<td>10/18</td>
<td>12/12</td>
<td>10/8</td>
<td>10/9</td>
</tr>
</tbody>
</table>

*Note.* TLE = temporal lobe epilepsy; HC = healthy controls; aMCI = amnestic mild cognitive impairment.
### Table 2

**Clinical and Demographic Characteristics of TLE Group**

<table>
<thead>
<tr>
<th>Pathology and Cognition</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onset Side (L/R/B)</td>
<td>11/11/6</td>
</tr>
<tr>
<td>HS+ (L/R/B)</td>
<td>11 (6/4/1)</td>
</tr>
<tr>
<td>CVLT (NI/I)</td>
<td>16 (9/7)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relevant Characteristics</th>
<th>Range</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education, years</td>
<td>10-18</td>
<td>12.9 (2.0)</td>
</tr>
<tr>
<td>Age of Seizure Onset, years</td>
<td>1-41</td>
<td>19.3 (11.5)</td>
</tr>
<tr>
<td>Disease Duration, years</td>
<td>1-47</td>
<td>14.2 (11.6)</td>
</tr>
</tbody>
</table>

*Note.* L = left, R = right, B = bilateral; HS+ = presence of hippocampal sclerosis; NI = not impaired, I = impaired.
### Table 3

**Percent Endorsed for Each Group and Each Stimulus and Response Type**

<table>
<thead>
<tr>
<th>Group</th>
<th>Target Old</th>
<th>Target Similar</th>
<th>Target New</th>
<th>Lure Old</th>
<th>Lure Similar</th>
<th>Lure New</th>
<th>Foil Old</th>
<th>Foil Similar</th>
<th>Foil New</th>
</tr>
</thead>
<tbody>
<tr>
<td>TLE</td>
<td>67.8</td>
<td>18.8</td>
<td>13.5</td>
<td>44.4</td>
<td>36.1</td>
<td>19.6</td>
<td>12.4</td>
<td>23.4</td>
<td>64.3</td>
</tr>
<tr>
<td></td>
<td>(18.0)</td>
<td>(12.5)</td>
<td>(13.0)</td>
<td>(15.6)</td>
<td>(17.2)</td>
<td>(13.8)</td>
<td>(13.3)</td>
<td>(13.9)</td>
<td>(18.3)</td>
</tr>
<tr>
<td>Young-HC</td>
<td>79.7</td>
<td>12.0</td>
<td>5.0</td>
<td>38.9</td>
<td>48.4</td>
<td>8.7</td>
<td>2.2</td>
<td>13.5</td>
<td>79.4</td>
</tr>
<tr>
<td></td>
<td>(12.1)</td>
<td>(6.8)</td>
<td>(4.8)</td>
<td>(12.1)</td>
<td>(15.3)</td>
<td>(8.8)</td>
<td>(2.6)</td>
<td>(6.8)</td>
<td>(9.7)</td>
</tr>
<tr>
<td>aMCI</td>
<td>75.6</td>
<td>10.8</td>
<td>13.5</td>
<td>58.5</td>
<td>19.1</td>
<td>22.4</td>
<td>12.4</td>
<td>10.4</td>
<td>77.1</td>
</tr>
<tr>
<td></td>
<td>(15.3)</td>
<td>(12.0)</td>
<td>(9.3)</td>
<td>(18.6)</td>
<td>(16.9)</td>
<td>(12.4)</td>
<td>(12.8)</td>
<td>(11.4)</td>
<td>(13.7)</td>
</tr>
<tr>
<td>Older-HC</td>
<td>82.6</td>
<td>6.5</td>
<td>4.5</td>
<td>50.1</td>
<td>31.4</td>
<td>9.6</td>
<td>3.1</td>
<td>9.2</td>
<td>76.8</td>
</tr>
<tr>
<td></td>
<td>(8.8)</td>
<td>(4.5)</td>
<td>(3.1)</td>
<td>(14.5)</td>
<td>(15.8)</td>
<td>(5.5)</td>
<td>(4.6)</td>
<td>(6.0)</td>
<td>(14.5)</td>
</tr>
</tbody>
</table>

*Note.* Percent endorsed for each group and each stimulus and response type (standard deviations below).
Figure Legends

Figure 1. The Mnemonic Similarity Task (A) The two phases of the MST showing examples of stimuli presented during the incidental encoding and subsequent “old”/“similar”/“new” recognition memory task. Colored boxes illustrate the three conditions, but were not shown during task administration. (B) Examples of images for each of the lure bins, ranging from most similar (lure bin 1) to least similar (lure bin 5). The figure was borrowed with the permission of the journal/authors and was previously published in Stark et al., 2019.

Figure 2. Memory Metrics of the Mnemonic Similarity Task. A) Lure discrimination index (LDI) for the TLE, aMCI, Young-HC, and Older-HC groups. B) Recognition memory discriminability scores for the four groups. The LDI responses are correct responses (“Similar” responses to lure items) minus incorrect responses (“Similar” responses to foil items). The recognition discriminability responses are hits (“Old” responses to target items) minus false alarms (“Old” responses to foils). TLE = temporal lobe epilepsy, young-HC = younger age-matched healthy controls, aMCI = amnestic mild cognitive impairment, older-HC = older age-matched healthy controls. Error bars indicate SEM. * group difference, p < 0.05, ** group difference, p < 0.01.

Figure 3. Behavioral Pattern Separation by Level of Similarity. A) Performance of TLE group and the young-HC and B) Performance of aMCI group and the older-HC as a function of the similarity between the target and the lure. Probability of correctly identifying foil items as “New” (i.e., avoiding the propensity to mistake them as old or previously seen) is shown. Of particular interest is performance with lure items, for which similarity with their corresponding targets is systematically varied in the MST as a way to differentially tax PS ability. Accordingly, performances shown here are binned by level of lure similarity, with items from lure bin 1 being
the most difficult to distinguish and items from lure bin 5 being the least difficult. C) Mean Z-scores depict the patient-versus-control differences shown in panels A (TLE) and B (aMCI). Individual patient performance was standardized according to the mean of the corresponding age-matched HC group, and these standardized (Z) scores are then averaged for each patient group. Abbreviations are as in Figure 2.

Figure 4. Temporal Lobe Epilepsy Performance on the MST and Clinical Measures. The dashed line represents the average performance from the age-matched healthy control group (young-HC) on each metric. A) Individuals with and without hippocampal sclerosis (HS+ and HS-), respectively. B) Individuals with left, right, or bilateral seizure onset. C) Individuals with or without impairment on CVLT-II long delay free recall, relative to age-adjusted norms. Impaired performance reflects a standard score 1.5 standard deviations or more below the age-based mean. Box plots show the minimum, first quartile, median, third quartile, and the maximum response values; outliers appear as black dots.
Figure 1

A) Encoding Phase
Indoor/Outdoor?

Test Phase
Old/Similar/New?

Repetition
Novel Foil
Similar Lure

B) Lure 1  Lure 2  Lure 3  Lure 4  Lure 5

Encoding

Test
Figure 2
Figure 3
Figure 4

A) Lure Discrimination Index

B) Seizure Laterality

C) Memory Impairment
### Supplementary Table 1

**Scores on Standardized Neuropsychological Testing for Healthy Controls (HC)**

<table>
<thead>
<tr>
<th>Neuropsychological test</th>
<th>Young-HC</th>
<th>Older-HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>29.3 (0.8)</td>
<td>29.1 (1.1)</td>
</tr>
<tr>
<td>RAVLT total</td>
<td>55.1 (8.1)</td>
<td>56.6 (7.9)</td>
</tr>
<tr>
<td>RAVLT immediate</td>
<td>11.7 (2.2)</td>
<td>13.2 (1.3)</td>
</tr>
<tr>
<td>RAVLT delay</td>
<td>12.2 (2.1)</td>
<td>13.0 (1.3)</td>
</tr>
<tr>
<td>Digit span</td>
<td>18.4 (3.6)</td>
<td>18.4 (4.3)</td>
</tr>
<tr>
<td>L-N sequencing</td>
<td>10.6 (2.3)</td>
<td>10.3 (2.0)</td>
</tr>
<tr>
<td>Trails A</td>
<td>20.0 (5.2)</td>
<td>26.8 (7.7)</td>
</tr>
<tr>
<td>Trails B</td>
<td>44.1 (14.1)</td>
<td>63.1 (15.9)</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td>44.0 (8.9)</td>
<td>48.1 (11.1)</td>
</tr>
<tr>
<td>Category fluency</td>
<td>22.7 (4.1)</td>
<td>21.4 (4.8)</td>
</tr>
<tr>
<td>WAIS III IQ</td>
<td>114.2 (7.5)</td>
<td>120.7 (9.7)</td>
</tr>
</tbody>
</table>

*Note. Mean scores (standard deviations). Mini Mental State Examination (MMSE), Rey Auditory Verbal Learning Test (RAVLT), Letter-Number Sequencing (L-N), Wechsler Adult Intelligence Scale 3rd Edition (WAIS III) Intelligence Quotient (IQ).*

*All scores presented are raw scores except for the WAIS III IQ which is a standardized IQ score with a mean of 100 and SD of 15.*
References


ketamine model suggests dentate gyrus pathology linked to N-methyl-D-aspartate receptor hypofunction. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(3), 231-238.


impairment in temporal lobe epilepsy with and without mesial temporal lobe sclerosis.

*Human Brain Mapping, 33, 489–499.*


*Human Brain Mapping, 31(9), 1339-1347.*


https://doi.org/10.1002/hipo.23299


https://doi.org/10.1212/wnl.54.3.581


