



# Evidence for a reactionary account of retrieval state initiation

Subin Han, Nicole M. Long

Department of Psychology, University of Virginia, Charlottesville, VA, United States

Corresponding Author: Nicole M. Long ([niclong@virginia.edu](mailto:niclong@virginia.edu))

## ABSTRACT

Engagement of the retrieval state (or mode) is theorized to be a precursor to successful retrieval, but precisely when the retrieval state is engaged is unclear. Our aim in the present study is to determine the time course of retrieval state initiation. Based on growing evidence that the retrieval state reflects internal attention, we hypothesize that the retrieval state is reactionary rather than preparatory, whereby the retrieval state is engaged following—rather than preceding—a stimulus. We collected scalp electroencephalography data during a mnemonic state task in which we explicitly biased participants to engage in either retrieval or encoding states, and in which we manipulated the stimulus onset asynchrony (SOA) between the instruction and stimulus onsets. Our general expectation is that selective engagement of the retrieval state will only occur after stimulus presentation and thus be unaffected by the SOA. Our behavioral results show that regardless of the SOA, a bias to retrieval compared with encoding leads to worse memory for object stimuli. Using cross-study multivariate pattern analyses, we find robust engagement of the retrieval state approximately 500 ms following stimulus onset and no evidence for selective retrieval state engagement during the instruction interval. Together, these findings suggest that retrieval state engagement is reactionary and is consistent with the retrieval state reflecting internal attention engaged in the attempt to access stored information.

**Keywords:** memory, encoding, retrieval, EEG, attention, MVPA

## 1. INTRODUCTION

Accessing stored episodic information underlies a host of behaviors and allows us to effectively engage with the world. For instance, as you plan dinner tonight, you may need to retrieve what you had for dinner last night. However, how we initiate this access process remains a critical open question. According to one view, a retrieval state (or “mode;” [Tulving, 1983](#)) supports episodic remembering whereby one intentionally initiates a retrieval state in response to an instruction to retrieve. Initiation of the retrieval state enables you to treat the stimulus—for example, “dinner”—as a retrieval cue in order to access stored episodic information. However, insofar as the retrieval state reflects internal attention ([Chun et al., 2011](#); [Long, 2023](#); [Long et al., 2018](#)) and a target is required for internal attention ([Chun et al., 2011](#)), recruitment of the retrieval state should only occur after,

not before, the stimulus appears. That is, it may be only after the target “dinner” is specified that you initiate the retrieval state, turning your mind’s eye inward to access what you had for dinner last night. Thus, whether we initiate retrieval prior to or following a stimulus remains an open question. The aim of this study was to determine the time course of retrieval initiation as the time course of retrieval is expected to impact downstream processing and behavioral outcomes.

The retrieval state may be preparatory, activated in response to an instruction to prepare to engage in episodic retrieval ([Rugg & Wilding, 2000](#)). Early work compared electrophysiological signals when participants were instructed to prepare to make a recognition (episodic) judgment vs. a semantic judgment ([Morcom & Rugg, 2002](#)). Approximately 600 ms after the instruction cue, a right frontal positivity emerges in response to the

Received: 20 February 2025 Revision: 10 October 2025 Accepted: 22 October 2025 Available Online: 4 November 2025



© 2025 The Authors. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

Imaging Neuroscience, Volume 3, 2025  
<https://doi.org/10.1162/IMAG.a.1023>

episodic vs. semantic instruction. Subsequent work has found a similar signal when contrasting instructions to prepare for episodic retrieval vs. a perceptual task in which no retrieval is required (Evans et al., 2015). The interpretation of these findings is that the retrieval state serves as a precursor to, but is separate from, episodic retrieval success. The individual first engages the retrieval state which enables them to treat the probe stimulus as an episodic retrieval cue. The retrieval state then allows the search process to begin which may or may not lead to access of the stored representation.

Recent work, however, suggests that the retrieval state may instead be reactionary. Using multivariate pattern analysis of spectral signals across the scalp, we have found that we can distinguish between retrieval and encoding states around 1000 ms following *stimulus* onset (Long & Kuhl, 2019; Smith et al., 2022). As a preparatory retrieval state would be engaged prior to the stimulus, we would have anticipated robust mnemonic state dissociations throughout the entirety of the stimulus interval. However, these late stimulus interval findings are in line with growing evidence from behavioral, neural, and computational investigations (Logan et al., 2021, 2024; Long, 2023) which suggest that the retrieval state may reflect internal attention. Whereas external attention is a focus on sensory, perceptual information, internal attention is the directing of the mind's eye inward to stored representations and task goals (Chun et al., 2011). In this framework, a target is needed upon which to focus internal attention which is consistent with mnemonic state dissociations following, rather than preceding, stimulus onset.

Whether retrieval state initiation is preparatory or reactionary is expected to have consequences for episodic remembering. Engaging the retrieval state both impacts how stimuli are processed (Long & Kuhl, 2021) and trades off with engaging the encoding state (Hasselman et al., 2002; Long & Kuhl, 2019) such that retrieving instead of encoding can negatively impact subsequent memory (Long & Kuhl, 2019). The time course of retrieval state initiation will clarify how episodic remembering unfolds and how probe stimuli are processed to support successful remembering. Insights into the timing of retrieval initiation will additionally inform our understanding of the relationship between mnemonic brain states and attentional processing.

Our hypothesis is that the retrieval state is reactionary, and sustained over time. Alternatively, the retrieval state may be preparatory—that is, engaged following the instruction to retrieve and prior to stimulus onset. To test our hypothesis, we conducted a mnemonic state task in which participants studied two sets of object images while recording scalp electroencephalography (EEG). The second set of object images categorically overlapped with the first set, and during the second set, participants

were explicitly instructed to either encode the second (present) object or retrieve the first (past) object. These instructions were intended to bias participants toward either an encoding or retrieval state. Our critical manipulation was the stimulus onset asynchrony (SOA) between the instruction and the stimulus onsets, whereby the shorter the SOA, the shorter the time between the instruction and the stimulus. Longer preparatory intervals can facilitate establishing task-appropriate cognitive states (Imburgio & Orr, 2021). Thus, we would expect that longer SOAs (i.e., increased preparation time) should facilitate participants' engagement in the retrieval state to the extent that it is preparatory. We used cross-study multivariate pattern analyses (MVPA) to measure retrieval state evidence (i.e., evidence for retrieval state engagement) over time. To the extent that the retrieval state is reactionary, we should find selective engagement of the retrieval state based on the instruction to encode or retrieve following stimulus onset for all SOA conditions.

## 2. MATERIALS AND METHODS

### 2.1. Participants

Forty (28 female; age range = 18–54 years, mean age = 25.6 years) English speakers from the University of Virginia community participated. This sample size is based on our prior work in which we find an effect of mnemonic state instructions on recognition memory (Smith et al., 2022) as described in the pre-registration report of this study (<https://osf.io/xnwg9>). All participants had normal or corrected-to-normal vision. Informed consent was obtained in accordance with the University of Virginia Institutional Review Board for Social and Behavioral Sciences, and participants were compensated for their participation. One participant was excluded from the final dataset due to technical issues resulting in poor signal quality throughout the majority of the session. Thus, data are reported for the remaining 39 participants.

### 2.2. Mnemonic state task experimental design

The task design follows our previous work (Smith et al., 2022) with one critical manipulation. Stimuli consisted of 864 object pictures, drawn from an image database with multiple exemplars per object category (Konkle et al., 2010). From this database, we chose 216 unique object categories and 4 exemplars from each category. For each participant, one exemplar in a set of four served as a List 1 object, one as a List 2 object, and the two remaining exemplars served as lures for the recognition phase. Object condition assignment was randomly generated for each participant.

### 2.2.1. General overview

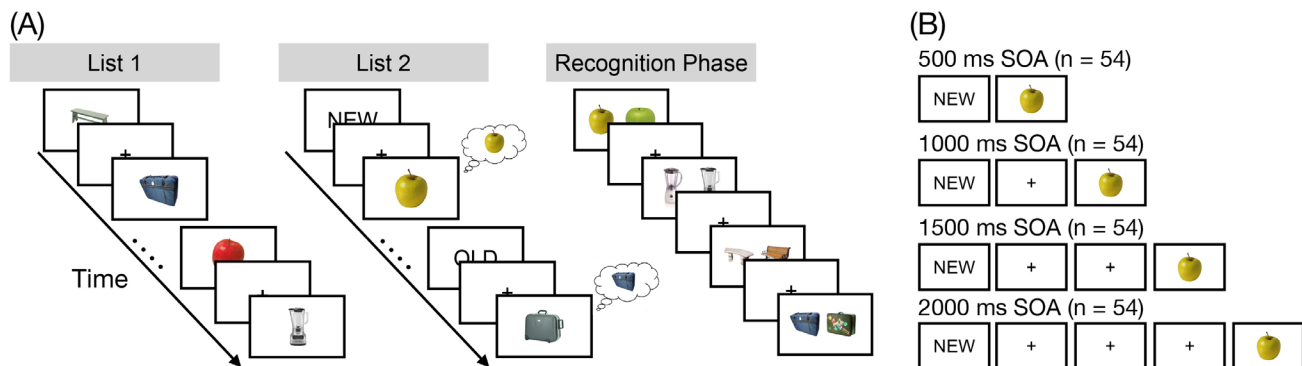
In each of 12 runs, participants viewed 2 lists, each containing 18 object images. For the first list (List 1), each object was new. For the second list (List 2), each object was again new but was categorically related to an object from the first list. For example, if List 1 contained an image of a bench, List 2 would contain an image of a different bench (Fig. 1A). During List 1, participants were instructed to encode each new object. During List 2, however, each trial contained an instruction to either encode the current object (e.g., the new bench) or to retrieve the corresponding object from List 1 (the old bench). The critical manipulation was the stimulus onset asynchrony (SOA) between the instruction and the stimulus onset. SOA ranged from 500 ms to 2000 ms in 500 ms increments (Fig. 1B) and varied across runs meaning that all List 2 trials within a run had the same duration. We held SOA constant within a run for two key reasons. First, a randomly changing SOA may induce carryover effects across trials, which could impact both behavior and retrieval state engagement. Second, a randomly changing SOA may reduce participants' ability to utilize the SOA to prepare for the upcoming stimulus. We have previously found both carryover effects and impacts on participants' mnemonic state engagement when we included a variable interstimulus interval between recognition trials (Wheelock & Long, 2024). SOA conditions were randomly assigned across runs for each participant. Following 12 runs, participants completed a

2-alternative forced-choice recognition test that separately assessed memory for all List 1 and List 2 objects.

**List 1.** On each trial, participants saw a single object presented for 2000 ms followed by a 1000 ms interstimulus interval. Participants were instructed to study the presented object in anticipation of a later memory test.

**List 2.** On each trial, participants saw a cue word, either "OLD" or "NEW" for 500 ms. Either immediately following the cue (SOA = 500 ms) or after a fixation cross (SOAs > 500 ms), an object image was presented for 2000 ms (Fig. 1B). All objects in List 2 were nonidentical exemplars drawn from the same category as the objects presented in the immediately preceding List 1. That is, if a participant saw a bench and an apple during List 1, a different bench and a different apple would be presented during List 2. On trials with a NEW instruction (encode trials), participants were to encode the presented object. On trials with an OLD instruction (retrieve trials), participants tried to retrieve the categorically related object from the preceding List 1. Importantly, this design prevented participants from completely ignoring List 2 objects following OLD instructions in that they could only identify the to-be-retrieved object category by processing the List 2 object.

To reduce potential confusion, we refer to trials with an OLD instruction as "retrieve" trials and trials with a NEW instruction as "encode" trials. Given that the terms "old" and "new" are traditionally used in the literature to denote



**Fig. 1.** Task design. (A) During List 1, participants studied individual objects (e.g., apple, suitcase). During List 2, participants saw novel objects that were from the same categories as the objects shown in List 1 (e.g., a new apple, a new suitcase). Preceding each List 2 object was an OLD instruction cue or a NEW instruction cue. The OLD cue signaled that participants were to retrieve the corresponding object from List 1 (e.g., the old suitcase). The NEW cue signaled that participants were to encode the current object (e.g., the new apple). Each run of the experiment contained a List 1 and List 2; object categories (e.g., apple) were not repeated across runs. During List 2 trials, we manipulated stimulus onset asynchrony (SOA) between the instruction onset and the stimulus onset. SOA varied across runs meaning that all List 2 trials within a run had the same SOA. After 12 runs, participants completed a 2-alternative forced-choice recognition test that tested memory for each List 1 and List 2 object. On each trial, a previously presented object, either from List 1 or List 2, was shown alongside a novel lure from the same category. Participants' task was to choose the previously presented object. List 1 and List 2 objects were never presented together. (B) There were four SOA conditions. The shortest SOA was 500 ms and the longest SOA was 2000 ms. The "NEW" or "OLD" instruction cue was presented for 500 ms. Either immediately following the cue (SOA = 500 ms) or after a delay of 500, 1000, or 1500 ms during which a fixation cross was shown (SOAs > 500 ms), an object image was presented for 2000 ms.

either probe status (e.g., targets are old, lures are new) or responses (participants' judgments of a probe's status), we prefer the condition labels "retrieve" and "encode" as they reflect the task that the participant was instructed to perform. However, we acknowledge that participants may not necessarily engage in either encoding or retrieval throughout the encode and retrieve trials, respectively.

Participants completed 12 runs with 2 lists in each run (List 1, List 2). Participants viewed 18 objects per list, yielding a total of 432 object stimuli from 216 unique object categories. Participants did not make a behavioral response during either List 1 or List 2. Following the 12 runs, participants completed a 2-alternative forced-choice recognition test.

### 2.2.2. Recognition phase

Following the 12 runs, participants completed the recognition phase. On each trial, participants saw two exemplars from the same object category (e.g., two benches; Fig. 1A). One object had previously been encountered either during List 1 or List 2. The other object was a lure and had not been presented during the experiment. Given that targets and lures were from the same object category, they shared many of the same perceptual features despite not being identical exemplars. Therefore, the strong overlap in shared physical features would make it more likely that participants relied on episodic or recollective-level information, rather than on familiarity or gist-level information to make their response (Brainerd & Reyna, 2002). Trials were self-paced, and participants selected (via button press) the previously presented object. Trials were separated by a 1000 ms interstimulus interval. There were a total of 432 recognition trials (corresponding to the 432 total List 1 and List 2 objects presented in the experiment). List 1 and List 2 objects never appeared in the same trial together; thus, participants never had to choose between two previously presented objects. Non-competitive testing of the objects enables us to separately evaluate the influence of instruction on List 1 and List 2 objects. List 1 and List 2 objects were presented randomly throughout the test phase.

### 2.3. EEG data acquisition and preprocessing

All acquisition and preprocessing methods are based on our previous work (Smith et al., 2022); for clarity we use the same text as previously reported. EEG recordings were collected using a BrainVision system and an Acti-Cap equipped with 64 Ag/AgCl active electrodes positioned according to the extended 10–20 system. All electrodes were digitized at a sampling rate of 1000 Hz and were referenced to electrode FCz. Offline, electrodes

were later converted to an average reference. Impedances of all electrodes were kept below 50 k $\Omega$ . Electrodes that demonstrated high impedance or poor contact with the scalp were excluded from the average reference; however, these electrodes were included in all subsequent analysis steps. Thus, no electrodes and no data are excluded for any participant. Bad electrodes were determined by voltage thresholding (see below).

Custom python codes were used to process the EEG data. We applied a high pass filter at 0.1 Hz, followed by a notch filter at 60 Hz and harmonics of 60 Hz to each participant's raw EEG data. We then performed three pre-processing steps (Nolan et al., 2010) to identify electrodes with severe artifacts. First, we calculated the mean correlation between each electrode and all other electrodes as electrodes should be moderately correlated with other electrodes due to volume conduction. We z-scored these means across electrodes and rejected electrodes with z-scores less than -3. Second, we calculated the variance for each electrode, as electrodes with very high or low variance across a session are likely dominated by noise or have poor contact with the scalp. We then z-scored variance across electrodes and rejected electrodes with a  $|z| \geq 3$ . Finally, we expect many electrical signals to be autocorrelated, but signals generated by the brain versus noise are likely to have different forms of autocorrelation. Therefore, we calculated the Hurst exponent, a measure of long-range autocorrelation, for each electrode and rejected electrodes with a  $|z| \geq 3$ . Electrodes marked as bad by this procedure ( $M = 1.85$ ,  $SD = 0.9097$ ) were excluded from the average re-reference. For 12 of the 39 participants, 1 or both of the electrodes used for eyeblink detection (Fp1, Fp2) were identified as bad and thus not included in the average re-referenced calculation. We then calculated the average voltage across all remaining electrodes at each time sample and re-referenced the data by subtracting the average voltage from the filtered EEG data. We used wavelet-enhanced independent component analysis (Castellanos & Makarov, 2006) to remove artifacts from eyeblinks and saccades.

### 2.4. EEG data analysis

We applied the Morlet wavelet transform (wave number 6) to the entire EEG time series across electrodes, for each of 46 logarithmically spaced frequencies (2–100 Hz; Smith et al., 2022). After log-transforming the power, we downsampled the data by taking a moving average across 100 ms time intervals from 4000 ms preceding to 4000 ms following object item presentation, and sliding the window every 25 ms, resulting in 317 time intervals (80 non-overlapping). Power values were then z-transformed by subtracting the mean and dividing by the stan-



dard deviation power. Mean and standard deviation power were calculated across all trials and across time points for each frequency.

## 2.5. Pattern classification analyses

Pattern classification analyses were performed using penalized (L2) logistic regression implemented via the sklearn module (0.24.2) in Python and custom Python code (Smith & Long, 2024). For all classification analyses, classifier features were composed of spectral power across 63 electrodes and 46 frequencies. Before pattern classification analyses were conducted, an additional round of z-scoring was performed across features (electrodes and frequencies) to eliminate trial-level differences in spectral power (Kuhl & Chun, 2014; Long & Kuhl, 2018; Smith et al., 2022). Therefore, mean univariate activity was matched precisely across all conditions and trial types. We extracted both “classification accuracy” and “classification evidence.” “Classification accuracy” represents a binary coding of whether the classifier successfully guessed the mnemonic label, encode or retrieve. We used classification accuracy for general assessment of classifier performance (i.e., whether encode/retrieve instructions could be decoded). “Classification evidence” is a continuous value reflecting the logit-transformed probability that the classifier assigns the correct mnemonic label (encode, retrieve) for each trial. Such an approach has been used to measure evidence for other forms of information, for example, faces vs. scenes (Lewis-Peacock & Norman, 2014) and faces vs. objects (DuBrow & Davachi, 2014). We used classification evidence as a trial-specific, continuous measure of mnemonic state information, which was used to assess the degree of retrieval evidence present on List 2 encode and retrieve trials. In effect, classification evidence indicates how confident the classifier is that a given pattern of spectral activity reflects either “encode” or “retrieve.”

### 2.5.1. Within-study mnemonic state classification

To demonstrate that participants selectively engage the retrieval state on retrieve trials, we conducted leave-one-run-out cross-validated classification (penalty parameter = 1) on the stimulus interval data collapsed across SOA. We trained each within-participant classifier to distinguish encode and retrieve trials based on z-scored spectral power averaged over the stimulus interval (2000 ms).

To test the extent to which the retrieval state is engaged during the instruction interval, we conducted leave-one-participant-out cross-validated classification (penalty parameter = 1). We trained each classifier (one per SOA condition, four in total) to distinguish encode and retrieve

trials based on z-scored spectral power averaged over the instruction interval (500 ms, 1000 ms, 1500 ms, or 2000 ms, respectively). We opted to average over time interval to run a single classifier for each SOA condition, rather than run multiple classifiers across multiple time intervals, which would necessitate different numbers of classifiers across SOA. We trained each classifier on the instruction-interval data from 38 participants, and tested on the instruction-interval data from the remaining 1 participant. We repeated this process iteratively for all 39 participants.

### 2.5.2. Cross-study mnemonic state classification

To measure retrieval state engagement for the List 2 instruction and stimulus intervals, we conducted three stages of classification using the same methods as in our prior work (Smith & Long, 2024). First, we conducted within-participant leave-one-run-out cross-validated classification (penalty parameter = 1) on all participants who completed the mnemonic state task ( $N = 103$ ; see Hong et al., 2023 for details). The classifier was trained to distinguish encoding versus retrieval states based on spectral power averaged across the 2000 ms stimulus interval during List 2 trials. For each participant, we generated true and null classification accuracy values. We permuted condition labels (encode and retrieve) for 1000 iterations to generate a null distribution for each participant. Any participant whose true classification accuracy fell above the 90th percentile of their respective null distribution was selected for further analysis ( $N = 37$ ; 28 female; age range = 18–42, mean age = 21.24 years). This thresholding reduces the degree to which noise contributes to the training data by excluding participants who do not have robust spectral dissociations between encode and retrieve trials. However, it is important to note that participants who do not meet this threshold may still selectively engage in encoding and retrieval states, but in a manner different from those who meet the threshold. Specifically, participants could switch between encoding and retrieval during a single trial which would reduce the ability of the classifier to correctly label said trial. Investigations of mnemonic state switching are outside of the scope of the present work, but may account for the variability in within-participant classification accuracy. Although we report findings based on the thresholded set of  $N = 37$  training participants, we find the same pattern of results when we use the full set of data ( $N = 103$ ) for training (see Supplementary Material). Next, we conducted leave-one-participant-out cross-validated classification (penalty parameter = 0.0001) on the selected participants to validate the mnemonic state classifier. We obtained classification accuracy of 60% which is significantly above chance ( $t_{36} = 6.000$ ,  $p < 0.0001$ ), indicating that the

cross-participant mnemonic state classifier is able to distinguish encoding and retrieval states. Finally, we applied the cross-participant mnemonic state classifier to the current data across both the stimulus and instruction intervals during List 2 trials, specifically in 100 ms intervals from 2000 ms preceding to 2000 ms following stimulus onset. We extracted classification evidence, the logit-transformed probability that the classifier assigned a given List 2 trial a label of encoding or retrieval. This approach provides a trial-level estimate of retrieval state evidence across the List 2 instruction and stimulus intervals.

## 2.6. Statistical analyses

We used repeated measures ANOVAs (rmANOVAs) and *t*-tests to assess the effect of instruction (encode, retrieve) and SOA on behavioral memory performance. We used rmANOVAs and *t*-tests to assess the effect of instruction, SOA, and time interval on retrieval state evidence during both the stimulus and instruction interval. We applied multiple comparisons corrections for post hoc tests (using False Discovery Rate correction).

We used paired-sample *t*-tests to compare classification accuracy across participants to chance-level decoding accuracy, as determined by permutation procedures. Namely, for each participant, we shuffled the condition labels of interest (e.g., encode and retrieve for the List 2 instruction classifier) and then calculated classification accuracy. We repeated this procedure 1000 times for each participant and then averaged the 1000 shuffled accuracy values for each participant. These mean values were used as participant-specific empirically derived measures of chance accuracy. We utilized the same general procedure to assess leave-one-participant-out classification accuracy.

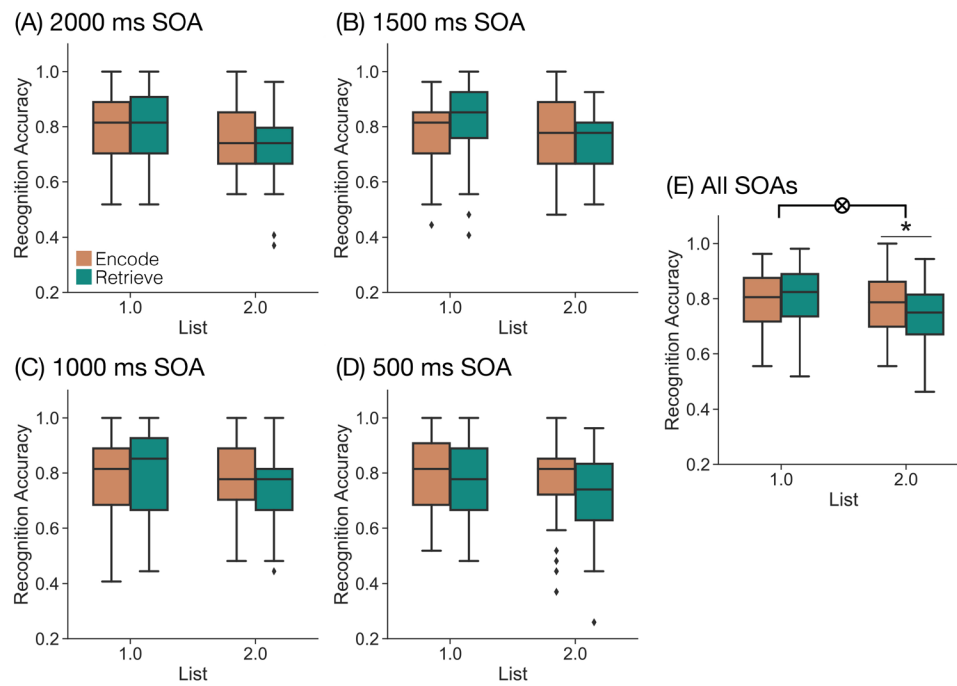
## 3. RESULTS

### 3.1. Memory accuracy is modulated by list and instruction, but not SOA

We first sought to replicate our prior finding that participants can shift between encoding and retrieval in a goal-directed manner (Long & Kuhl, 2019; Smith et al., 2022). We expected to find higher accuracy for List 2 items studied with the encode relative to the retrieve instruction given the assumption that encoding and retrieval trade off. We assessed recognition accuracy specifically for the 2000 ms SOA condition (Fig. 2A) as this matches our prior work. Following our pre-registration, we conducted a  $2 \times 2$  rmANOVA with factors of List (1, 2) and instruction (encode, retrieve) and recognition accuracy as the dependent variable. We found a significant main effect of list ( $F_{1,38} = 16.07$ ,  $p = 0.0003$ ,  $\eta_p^2 = 0.2972$ ) driven by greater recognition

accuracy for List 1 items ( $M = 0.7963$ ,  $SD = 0.1097$ ) compared with List 2 items ( $M = 0.7407$ ,  $SD = 0.1103$ ). We did not find a significant main effect of instruction ( $F_{1,38} = 1.104$ ,  $p = 0.3001$ ,  $\eta_p^2 = 0.0282$ ). We did not find a significant interaction between list and instruction ( $F_{1,38} = 2.985$ ,  $p = 0.0922$ ,  $\eta_p^2 = 0.0728$ ). Although we failed to find a significant list by instruction interaction as in our prior work, we expect that this is due to the reduced number of trials in the present analysis. With the SOA manipulation, the number of trials per condition (list crossed with instruction) is 27 (3 runs, 18 trials per list per run, half encode/half retrieve) which is substantially lower than the trial count in our prior work (72). The List 2 dissociation, although not significant ( $t_{38} = 1.980$ ,  $p = 0.055$ ,  $d = 0.2908$ ), follows our previous observation and current prediction whereby memory accuracy is numerically greater for encode ( $M = 0.7588$ ,  $SD = 0.1219$ ) relative to retrieve trials ( $M = 0.7227$ ,  $SD = 0.1263$ ).

Our hypothesis is that the retrieval state is reactionary, or engaged following stimulus onset. Thus, SOA should not modulate the list by instruction interaction. That is, participants should engage in encoding on trials with the encode instruction and in retrieval on trials with the retrieve instruction after the stimulus onset regardless of the length of the instruction–stimulus interval. However, to the extent that the retrieval state is preparatory, with a longer SOA, participants will have more time to selectively engage in each mnemonic state prior to stimulus onset. Thus, with longer SOAs, participants should have greater engagement in the respective mnemonic states which will result in a larger interaction between list and instruction on memory accuracy as SOA increases. To test these two alternatives and following our pre-registration, we conducted a  $2 \times 2 \times 4$  rmANOVA with factors of List (1, 2), instruction (encode, retrieve), and SOA (500 ms, 1000 ms, 1500 ms, 2000 ms) and recognition accuracy as the dependent variable (Fig. 2). We found a significant main effect of list ( $F_{1,38} = 15.31$ ,  $p = 0.0004$ ,  $\eta_p^2 = 0.2872$ ) driven by greater recognition accuracy for List 1 items ( $M = 0.7926$ ,  $SD = 0.1145$ ) than for List 2 items ( $M = 0.7518$ ,  $SD = 0.1049$ ). We found a significant main effect of instruction ( $F_{1,38} = 5.113$ ,  $p = 0.0296$ ,  $\eta_p^2 = 0.1186$ ) driven by greater recognition accuracy for encode ( $M = 0.7787$ ,  $SD = 0.1046$ ) than for retrieve trials ( $M = 0.7657$ ,  $SD = 0.1081$ ). We did not find a significant main effect of SOA ( $F_{3,114} = 0.6709$ ,  $p = 0.5716$ ,  $\eta_p^2 = 0.0174$ ). As we expected, we found a significant interaction between list and instruction ( $F_{1,38} = 12.52$ ,  $p = 0.0011$ ,  $\eta_p^2 = 0.2478$ ). This interaction was driven by significantly greater recognition accuracy for List 2 encode ( $M = 0.7707$ ,  $SD = 0.11$ ) than for List 2 retrieve trials ( $M = 0.7329$ ,  $SD = 0.1074$ ;  $t_{38} = 4.109$ ,  $p = 0.0002$ ,  $d = 0.3472$ , FDR corrected), and numerically greater recognition accuracy for List 1 retrieve ( $M = 0.7984$ ,  $SD = 0.1232$ ) than for



**Fig. 2.** Influence of mnemonic instructions on memory behavior. We assessed recognition accuracy as a function of List (1, 2) and instruction (orange, encode; teal, retrieve) separately for each SOA condition (A–D) and for all trials averaged across SOA conditions (E). Box-and-whisker plots show the interquartile range, with median indicated by the horizontal line and whiskers denoting maximum and minimum (mean  $\pm$  2.5  $\times$  standard deviation). Diamonds indicate outliers. (A–D) We did not find any effect of SOA on recognition accuracy. (E) We found a significant interaction between list and instruction ( $p = 0.0011$ ) driven by greater accuracy for List 2 encode than for List 2 retrieve trials.  $*p < 0.05$ , FDR corrected.

List 1 encode trials ( $M = 0.7868$ ,  $SD = 0.1121$ ;  $t_{38} = -1.3032$ ,  $p = 0.2003$ ,  $d = -0.0988$ ). We did not find a significant interaction between either list and SOA ( $F_{3,114} = 0.6241$ ,  $p = 0.6009$ ,  $\eta_p^2 = 0.0162$ ) or instruction and SOA ( $F_{3,114} = 0.7399$ ,  $p = 0.5304$ ,  $\eta_p^2 = 0.0191$ ). We did not find a significant three-way interaction between list, instruction, and SOA ( $F_{3,114} = 0.3737$ ,  $p = 0.7721$ ,  $\eta_p^2 = 0.0097$ ). Bayes factor analysis revealed that a model without the three-way interaction term ( $H_0$ ) is preferred to a model with the three-way interaction term ( $H_1$ ;  $H_{10} = 0.0146$  very strong evidence for  $H_0$ ).

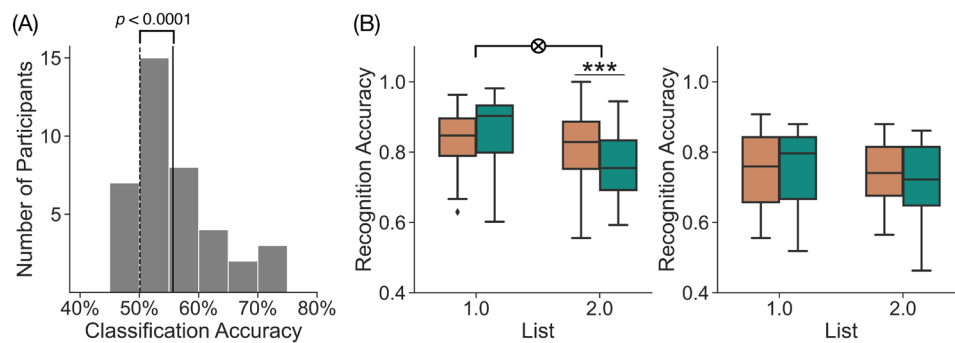
Together, our behavioral results support the reactionary account. We find robust evidence that participants can selectively engage in memory encoding and memory retrieval, replicating our past work. We do not find evidence that the amount of preparation time or SOA between instruction and stimulus influences the impact of list and instruction on memory accuracy.

### 3.2. Participants selectively engage the retrieval state

To demonstrate that participants selectively engage the retrieval state on retrieve trials, we performed within-participant, leave-one-run-out cross-validated mnemonic state decoding on z-scored power averaged

across the stimulus interval (0 to 2000 ms). We found significantly above-chance mean classification accuracy ( $M = 55.67\%$ ,  $SD = 6.925\%$ ,  $t_{38} = 4.927$ ,  $p < 0.0001$ ,  $d = 1.142$ ; Fig. 3A). Eighteen of 39 participants (46%) had true classification accuracy values that exceeded the 90th percentile of their chance distribution. Thus, replicating our past findings (Long & Kuhl, 2019; Smith et al., 2022), we found that the spectral patterns associated with encode and retrieve trials are distinguishable, suggesting that participants selectively engage mnemonic states in response to the instructions.

The extent to which participants selectively engage in mnemonic states should be related to behavioral performance on the two-alternative forced-choice recognition task. That is, participants who engage the retrieval state selectively on retrieve trials should show a stronger list by interaction dissociation than participants who do not. To link classification performance with behavior, we conducted an exploratory analysis of recognition accuracy separately for those participants who demonstrated above- vs. below-chance decoding accuracy (Fig. 3B). For the 18 above-chance participants, we found a significant list by instruction interaction ( $F_{1,17} = 13.06$ ,  $p = 0.0021$ ,  $\eta_p^2 = 0.4344$ ) driven by significantly greater recognition accuracy for List 2 encode ( $M = 0.8117$ ,  $SD = 0.1106$ ) than for List 2 retrieve ( $M = 0.7603$ ,



**Fig. 3.** Stimulus-interval classification accuracy and influence of mnemonic instructions on memory behavior.

(A) We performed leave-one-run-out within-participant classification on the stimulus interval and measured classification accuracy for each participant. Mean classification accuracy across all participants (solid vertical line) is shown along with a histogram of classification accuracies for individual participants (gray bars) and mean classification accuracy for permuted data across all participants (dashed vertical line). Mean classification accuracy for permuted data ranged from 49.37% to 50.79% across individual participants (1000 permutations per participant). Mean classification accuracy was 55.67%, which differed significantly from chance ( $p < 0.0001$ ). (B) We assessed recognition accuracy as a function of List (1, 2) and instruction (orange, encode; teal, retrieve) separately for above-chance (left) and below-chance (right) participants. For the 18 above-chance participants, we found a significant interaction between list and instruction ( $p = 0.0008$ ) driven by greater accuracy for List 2 encode than for List 2 retrieve trials. For the 21 below-chance participants, we did not find a significant list by instruction interaction ( $p = 0.1489$ ). \*\*\* $p < 0.001$ .

$SD = 0.0932$ ) trials ( $t_{17} = 4.076$ ,  $p = 0.0008$ ,  $d = 0.503$ ) and marginally greater recognition accuracy for List 1 retrieve ( $M = 0.8544$ ,  $SD = 0.1098$ ) than for List 1 encode ( $M = 0.8292$ ,  $SD = 0.0928$ ) trials ( $t_{17} = 2.02$ ,  $p = 0.0594$ ,  $d = 0.248$ ). In contrast, the 21 below-chance participants did not show a significant list by instruction interaction ( $F_{1,20} = 2.253$ ,  $p = 0.1489$ ,  $\eta_p^2 = 0.1013$ ). A direct between-group comparison shows a marginally significant interaction ( $F_{1,37} = 3.485$ ,  $p = 0.0699$ ,  $\eta_p^2 = 0.09$ ), suggesting that selective retrieval state engagement is linked to subsequent recognition accuracy.

### 3.3. Retrieval state evidence is modulated by instruction after stimulus onset, but not before

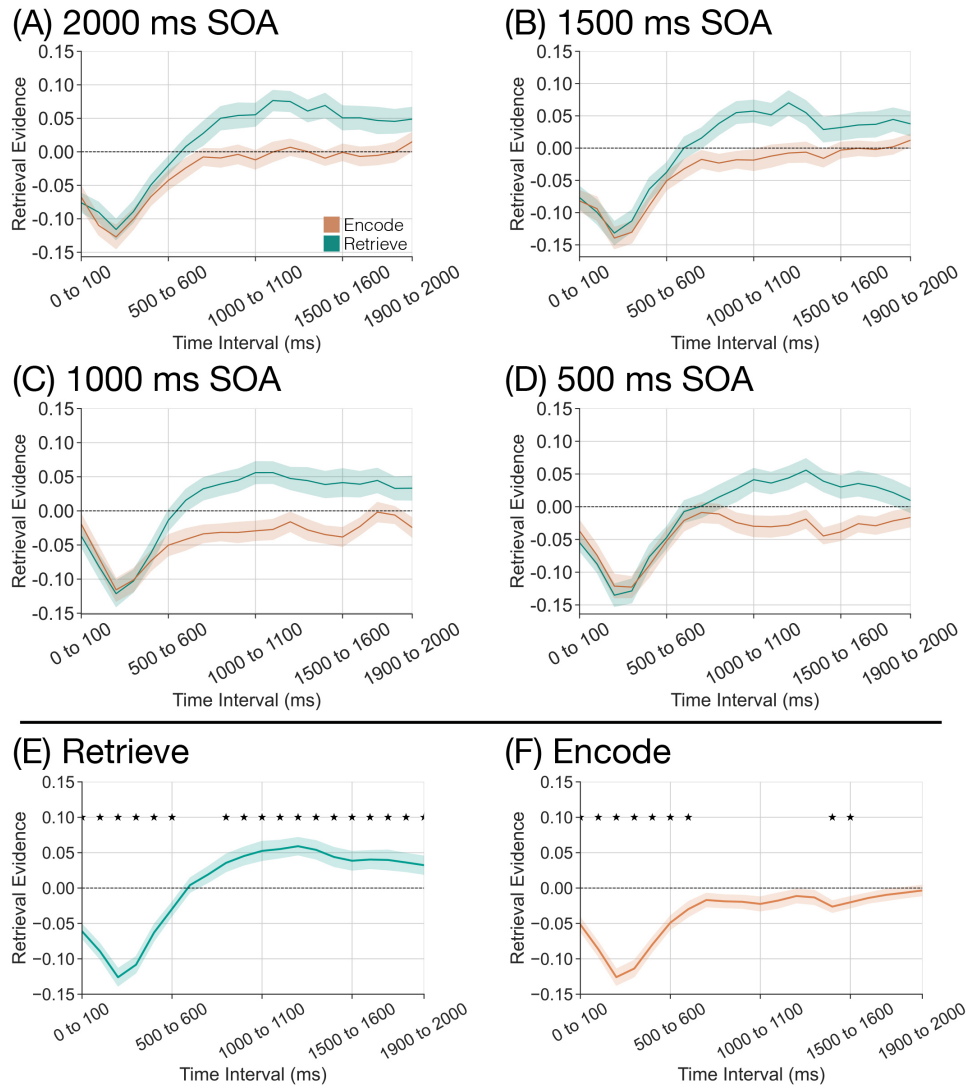
Our central goal was to test the hypothesis that the retrieval state is reactionary, and sustained over time. To test our hypothesis, we applied a cross-study mnemonic state classifier to trial-level spectral data across the whole brain. We leveraged the high temporal resolution of EEG and examined retrieval state evidence for 100 ms time intervals during the instruction and stimulus intervals. Specifically, we extracted retrieval state evidence from 500 ms to 2000 ms preceding (based on the SOA condition) to 2000 ms following the stimulus onset. To the extent that the retrieval state is reactionary (i.e., engaged after stimulus onset), we should find no dissociation in retrieval state evidence on the basis of instruction during the instruction interval and only find a dissociation during the stimulus interval. However, to the extent that the retrieval state is preparatory (i.e., engaged in response to

the instruction cue), we should find a dissociation in retrieval state evidence on the basis of instruction during the instruction interval. There should be greater dissociation with longer SOAs, and the dissociation should linger into the stimulus interval.

We first sought to replicate our prior finding that there is greater retrieval evidence on retrieve than on encode trials later in the stimulus interval (Smith et al., 2022). For this analysis, we focused exclusively on the 2000 ms SOA condition given that we used a fixed 2000 ms SOA in our prior work.

Following our pre-registration, we conducted a  $2 \times 20$  rmANOVA with factors of instruction (encode, retrieve) and time interval (twenty 100 ms time intervals across the 2000 ms stimulus interval) and retrieval state evidence as the dependent variable (Fig. 4A). We found a significant main effect of instruction ( $F_{1,38} = 10.56$ ,  $p = 0.0024$ ,  $\eta_p^2 = 0.2174$ ) driven by greater retrieval evidence for retrieve ( $M = 0.0139$ ,  $SD = 0.0621$ ) than for encode trials ( $M = -0.0286$ ,  $SD = 0.0555$ ). We found a significant main effect of time interval ( $F_{19,722} = 28.37$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.4274$ ). We found a significant interaction between instruction and time interval ( $F_{19,722} = 2.392$ ,  $p = 0.0008$ ,  $\eta_p^2 = 0.0592$ ) driven by greater retrieval evidence for retrieve than for encode trials later in the stimulus interval (Fig. 4A). Thus, we successfully replicated our prior finding that retrieval state evidence differs as a function of instruction during the stimulus interval. Furthermore, these findings suggest that similar spectral patterns are recruited to support retrieval throughout the latter half of the stimulus interval. If different spectral patterns were





**Fig. 4.** Influence of mnemonic instructions on stimulus-interval retrieval state evidence. We applied a cross-study classifier to the stimulus interval (twenty 100 ms time intervals) to measure retrieval state evidence over time. Error bars reflect standard error of the mean. (A–D) We assessed retrieval state evidence as a function of instruction (encode, orange; retrieve, teal) separately for each stimulus onset asynchrony (SOA) condition. Across all SOA conditions, we find a dissociation in retrieval state evidence on the basis of instruction around 500 ms after stimulus onset. (E–F) We averaged across SOA conditions separately for retrieve and encode trials, and performed 20 one-sample t-tests against zero.  $*p < 0.05$ , FDR-corrected.

active during different time intervals, the cross-study classifier would yield evidence values around 0. Instead, the classifier demonstrates consistent high confidence (positive evidence values) that the spectral activity patterns on retrieve trials appear retrieval like.

Next, we evaluated the effect of SOA on retrieval state evidence during the stimulus interval. To the extent that the retrieval state is reactionary, we should find no effect of SOA on retrieval state evidence. However, to the extent that the retrieval state is preparatory, there should be a greater instruction-interval dissociation in retrieval state evidence as a function of instruction as SOA increases. Thus, we should find an effect of SOA during the stimulus interval whereby there is a greater dissocia-

tion in retrieval state evidence for shorter SOA conditions than for longer SOA conditions. Following our pre-registration, we conducted a  $2 \times 4 \times 20$  rmANOVA with factors of instruction (encode, retrieve), SOA (500 ms, 1000 ms, 1500 ms, 2000 ms), and time interval (twenty 100 ms time intervals) and retrieval state evidence as the dependent variable (Fig. 4A–D). We found a significant main effect of instruction ( $F_{1,38} = 15.23$ ,  $p = 0.0004$ ,  $\eta_p^2 = 0.2861$ ) driven by greater retrieval evidence for retrieve ( $M = 0.0039$ ,  $SD = 0.0537$ ) than for encode trials ( $M = -0.0368$ ,  $SD = 0.04$ ). We found a significant main effect of time interval ( $F_{19,722} = 45.30$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.5438$ ). We did not find a significant main effect of SOA ( $F_{3,114} = 1.217$ ,  $p = 0.3068$ ,  $\eta_p^2 = 0.031$ ).

We found a significant interaction between instruction and time interval ( $F_{19,722} = 9.9577, p < 0.0001, \eta_p^2 = 0.2076$ ), and a significant interaction between SOA and time interval ( $F_{57,2166} = 1.602, p = 0.0031, \eta_p^2 = 0.0405$ ). We did not find a significant interaction between instruction and SOA ( $F_{3,114} = 0.2892, p = 0.8331, \eta_p^2 = 0.0076$ ). We did not find a significant three-way interaction between instruction, SOA, and time interval ( $F_{57,2166} = 0.6754, p = 0.9702, \eta_p^2 = 0.0175$ ). Bayes factor analysis revealed that a model without the three-way interaction term ( $H_0$ ) is preferred to a model with the three-way interaction term ( $H_1$ ;  $H_{10} = 1.982 \times 10^{-8}$  extreme evidence for  $H_0$ ). In short, we found no effect of SOA on retrieval state evidence during the stimulus interval which supports the reactionary account whereby mnemonic states are engaged following stimulus onset.

As in our prior work (e.g., Long & Kuhl, 2019), we find a significant dissociation in retrieval state evidence as a function of instruction. To specifically test the degree to which the retrieval state is engaged on retrieve and encode trials, we performed 20 one-sample  $t$ -tests (twenty 100 ms time intervals) averaged across SOA conditions separately for each instruction (Fig. 4E-F). Due to the structure of the independent classifier, significant positive values suggest engagement of the retrieval state, whereas significant negative values suggest engagement of the encoding state. We report the results in Table 1 and summarize the key findings here. We found significant decreases in retrieval evidence for both retrieve and

encode trials for the first ~500 ms following stimulus onset. For retrieve trials, we found a significant increase in retrieval evidence for all subsequent time intervals starting at 800 to 900 ms following stimulus onset. For encode trials, we found a significant decrease in retrieval evidence from 1400 ms to 1600 ms. Thus, regardless of instruction, participants appear to initially engage the encoding state, possibly to support perception of the List 2 item. This initial encoding state engagement is followed by robust, sustained engagement of the retrieval state on retrieve trials, with more transient engagement of the encoding state later on encode trials.

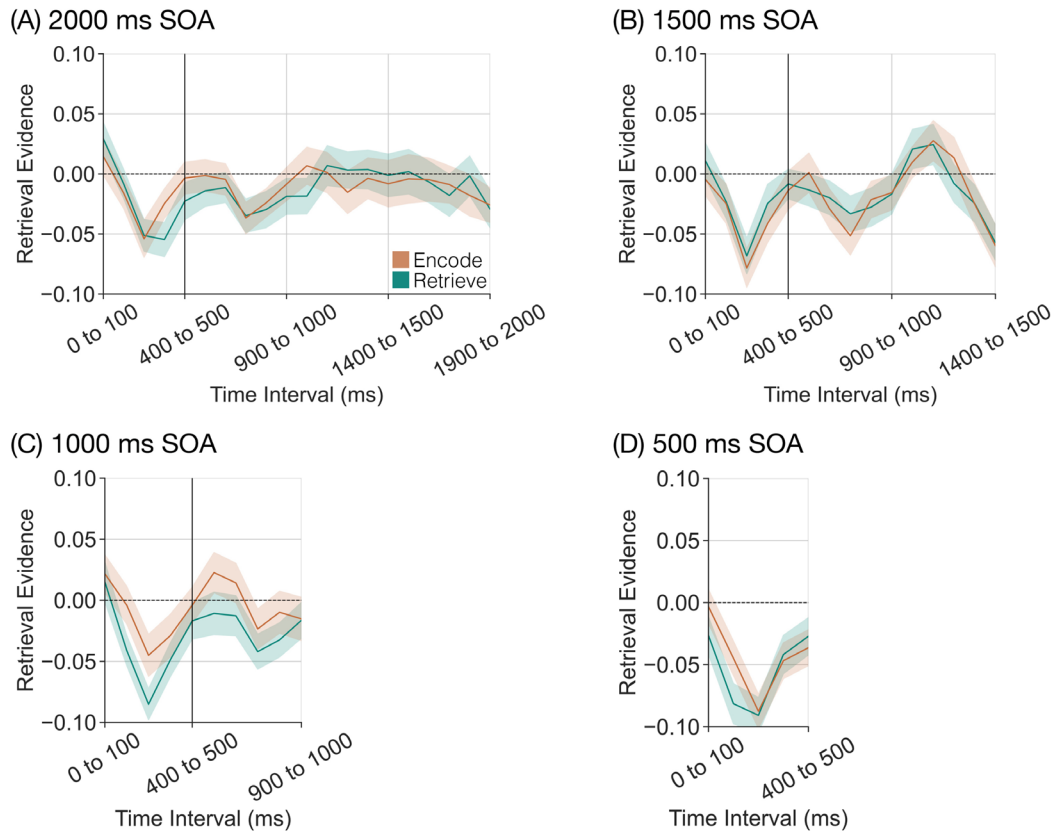
To investigate the extent to which the retrieval state is recruited during the instruction interval, we evaluated the effect of instruction on instruction-interval retrieval state evidence separately for each SOA condition (Fig. 5). To the extent that the retrieval state is preparatory, we should find greater retrieval state evidence on retrieve than on encode trials during the instruction interval. Following our pre-registration, we conducted four separate rmANOVAs. Each ANOVA had factors of instruction (encode, retrieve) and time interval (from five to twenty 100 ms time intervals depending on the SOA condition) and retrieval state evidence as the dependent variable. We report the results of this ANOVA in Table 2 and highlight the key findings here.

Except for the 1000 ms SOA condition, we did not find a significant main effect of instruction, meaning that retrieval evidence did not significantly differ across

**Table 1.** Comparison of retrieval state evidence vs. zero via one-sample  $t$ -tests.

| Time interval   | Retrieve trials |                 |        | Encode trials |                 |        |
|-----------------|-----------------|-----------------|--------|---------------|-----------------|--------|
|                 | $t_{38}$        | $p$             | $d$    | $t_{38}$      | $p$             | $d$    |
| 0 to 100 ms     | -6.01           | < <b>0.0001</b> | 0.975  | -5.131        | < <b>0.0001</b> | 0.8324 |
| 100 to 200 ms   | -8.031          | < <b>0.0001</b> | 1.303  | -8.374        | < <b>0.0001</b> | 1.359  |
| 200 to 300 ms   | -10.04          | < <b>0.0001</b> | 1.629  | -10.89        | < <b>0.0001</b> | 1.767  |
| 300 to 400 ms   | -9.228          | < <b>0.0001</b> | 1.497  | -9.645        | < <b>0.0001</b> | 1.565  |
| 400 to 500 ms   | -5.248          | < <b>0.0001</b> | 0.8513 | -7.756        | < <b>0.0001</b> | 1.258  |
| 500 to 600 ms   | -2.974          | <b>0.0051</b>   | 0.4824 | -5.0          | < <b>0.0001</b> | 0.8111 |
| 600 to 700 ms   | 0.4054          | 0.6875          | 0.0658 | -2.749        | <b>0.0091</b>   | 0.446  |
| 700 to 800 ms   | 1.656           | 0.1059          | 0.2687 | -1.739        | 0.0902          | 0.282  |
| 800 to 900 ms   | 2.837           | <b>0.0072</b>   | 0.4602 | -2.026        | 0.0499          | 0.3286 |
| 900 to 1000 ms  | 3.538           | <b>0.0011</b>   | 0.5739 | -2.063        | 0.0460          | 0.3347 |
| 1000 to 1100 ms | 3.838           | <b>0.0005</b>   | 0.6225 | -2.223        | 0.0322          | 0.3607 |
| 1100 to 1200 ms | 4.367           | < <b>0.0001</b> | 0.7085 | -1.628        | 0.1117          | 0.2642 |
| 1200 to 1300 ms | 4.743           | < <b>0.0001</b> | 0.7694 | -1.145        | 0.2593          | 0.1858 |
| 1300 to 1400 ms | 4.098           | <b>0.0002</b>   | 0.6647 | -1.298        | 0.2020          | 0.2106 |
| 1400 to 1500 ms | 3.227           | <b>0.0026</b>   | 0.5235 | -3.121        | <b>0.0034</b>   | 0.5062 |
| 1500 to 1600 ms | 2.873           | <b>0.0066</b>   | 0.4661 | -2.458        | <b>0.0186</b>   | 0.3988 |
| 1600 to 1700 ms | 3.14            | <b>0.0033</b>   | 0.5094 | -1.869        | 0.0694          | 0.3031 |
| 1700 to 1800 ms | 2.808           | <b>0.0078</b>   | 0.4554 | -1.165        | 0.2515          | 0.1889 |
| 1800 to 1900 ms | 2.769           | <b>0.0086</b>   | 0.4492 | -0.8719       | 0.3887          | 0.1414 |
| 1900 to 2000 ms | 2.473           | <b>0.018</b>    | 0.4011 | -0.4509       | 0.6546          | 0.0731 |

Bold values indicate  $p < 0.05$ , FDR-corrected.



**Fig. 5.** Influence of mnemonic instructions on instruction-interval retrieval state evidence. We applied a cross-study classifier to the instruction interval (five to twenty 100 ms time intervals depending on the SOA condition) to measure retrieval state evidence as a function of instruction (encode, orange; retrieve, teal) separately for each stimulus onset asynchrony (SOA) condition (A–D). The vertical line at the 400 to 500 ms time interval indicates the offset of the instruction cue. Except for the 1000 ms SOA condition, we find no dissociation in retrieval evidence on the basis of instruction. (C) In the 1000 ms SOA condition, we find greater retrieval evidence for encode than for retrieve trials ( $p = 0.0325$ ). Error bars reflect standard error of the mean.

retrieve vs. encode instructions. Bayes factor analysis revealed that a model without instruction ( $H_0$ ) is preferred to a model with instruction for the three SOA conditions (500 ms SOA:  $H_1$ ;  $H_{10} = 0.0263$  very strong evidence for  $H_0$ ; 1500 ms SOA:  $H_1$ ;  $H_{10} = 2.730 \times 10^{-5}$  extreme evidence for  $H_0$ ; 2000 ms SOA:  $H_1$ ;  $H_{10} = 1.204 \times 10^{-5}$  extreme evidence for  $H_0$ ). Thus, we do not find evidence that participants selectively engage the retrieval state during the instruction interval. The significant effect that we found for the 1000 ms SOA condition was in the opposite direction as we predicted—retrieval state evidence was greater for encode ( $M = -0.0072$ ,  $SD = 0.077$ ) relative to retrieve trials ( $M = -0.0291$ ,  $SD = 0.0604$ ). Thus, we failed to find evidence to support the interpretation that the retrieval state is preparatory.

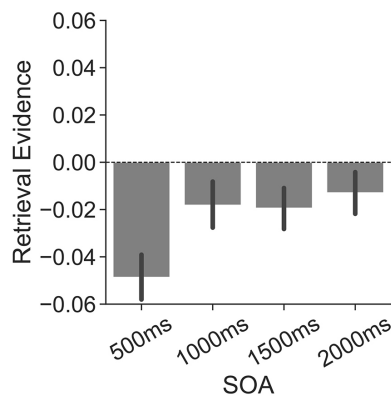
Finally, we directly compared instruction-interval retrieval state evidence across SOA conditions. Based on prior work demonstrating that participants can engage control processes in preparation for an upcoming event (Gruber & Otten, 2010), we hypothesized that mnemonic

state evidence might be modulated by SOA. Our classifier is designed such that less evidence for the retrieval state (negative evidence values) is equivalent to greater evidence for the encoding state. Insofar as the encoding state is engaged in anticipation of the upcoming stimulus, we expected that encoding state evidence would increase—or, conversely, that retrieval state evidence would decrease—as the SOA increased. Following our pre-registration, we averaged retrieval state evidence across the instruction interval to account for the variable instruction interval length. We conducted a  $2 \times 4$  rmANOVA with factors of instruction (encode, retrieve) and SOA (500 ms, 1000 ms, 1500 ms, 2000 ms), and retrieval state evidence as the dependent variable. We found a significant main effect of SOA ( $F_{3,114} = 3.957$ ,  $p = 0.0100$ ,  $\eta_p^2 = 0.0943$ ). However, counter to our expectations, we found greater retrieval evidence for longer SOA than for shorter SOA conditions (Fig. 6). The main effect of instruction and the interaction between instruction and SOA were not significant (main effect:

**Table 2.** ANOVA results of the effect of instruction and time interval on instruction-interval retrieval state evidence separately for each SOA condition.

|            | Main effect of instruction | Main effect of time | Interaction of instruction × time |
|------------|----------------------------|---------------------|-----------------------------------|
|            |                            | 500 ms              |                                   |
| df         | (1,38)                     | (4,152)             | (4,152)                           |
| F          | 0.5128                     | 13.477              | 2.033                             |
| p          | 0.4783                     | < <b>0.0001</b>     | 0.0925                            |
| $\eta_p^2$ | 0.0133                     | 0.2618              | 0.0508                            |
|            |                            | 1000 ms             |                                   |
| df         | (1,38)                     | (9,342)             | (9,342)                           |
| F          | 4.924                      | 6.943               | 0.7532                            |
| p          | <b>0.0325</b>              | < <b>0.0001</b>     | 0.6599                            |
| $\eta_p^2$ | 0.1147                     | 0.1545              | 0.0194                            |
|            |                            | 1500 ms             |                                   |
| df         | (1,38)                     | (14,532)            | (14,532)                          |
| F          | 0.0794                     | 8.050               | 0.5036                            |
| p          | 0.7797                     | < <b>0.0001</b>     | 0.9315                            |
| $\eta_p^2$ | 0.0021                     | 0.1748              | 0.0131                            |
|            |                            | 2000 ms             |                                   |
| df         | (1,38)                     | (19,722)            | (19,722)                          |
| F          | 0.0473                     | 3.085               | 0.799                             |
| p          | 0.8291                     | < <b>0.0001</b>     | 0.7098                            |
| $\eta_p^2$ | 0.0012                     | 0.0751              | 0.0206                            |

Bold values indicate  $p < 0.05$ .

**Fig. 6.** Influence of SOA on instruction-interval retrieval state evidence. We averaged retrieval state evidence across the instruction interval (five to twenty 100 ms time intervals depending on the SOA condition) and instruction (encode, retrieve) for each SOA condition. We find less retrieval state evidence for shorter SOA conditions. Error bars reflect standard error of the mean.

$F_{1,38} = 1.730$ ,  $p = 0.1963$ ,  $\eta_p^2 = 0.0435$ ; interaction:  $F_{3,114} = 1.072$ ,  $p = 0.3639$ ,  $\eta_p^2 = 0.0274$ .

We directly compared retrieval state evidence for each SOA condition with zero. Specifically, we performed one-sample  $t$ -tests on retrieval evidence averaged across the instruction interval. For all SOA conditions, we found negative retrieval evidence, although this effect was only significant for the 500 ms SOA condition following FDR correction (500 ms:  $M = -0.0487$ ,  $SD = 0.0566$ ,

$t_{38} = -5.299$ ,  $p < 0.0001$ ,  $d = 0.8596$ ; 1000 ms:  $M = -0.0181$ ,  $SD = 0.0612$ ,  $t_{38} = -1.829$ ,  $p = 0.0752$ ,  $d = 0.2967$ ; 1500 ms:  $M = -0.0194$ ,  $SD = 0.053$ ,  $t_{38} = -2.256$ ,  $p = 0.0299$ ,  $d = 0.366$ ; 2000 ms:  $M = -0.0129$ ,  $SD = 0.056$ ,  $t_{38} = -1.422$ ,  $p = 0.1633$ ,  $d = 0.2306$ ). These findings suggest that participants engage the encoding state during the instruction interval and to a greater degree for shorter SOAs.

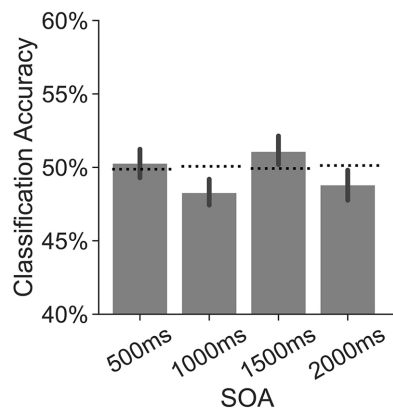
### 3.4. Instruction-interval decoding does not exceed chance

In the preceding analyses, we used an independent mnemonic state classifier to estimate retrieval state evidence. This approach was motivated by both technical and conceptual considerations. First, due to the SOA manipulation, the number of trials available for within-participant mnemonic state decoding during the instruction interval was relatively low ( $n = 54$  total trials, 27 per class). Thus, a failure to decode mnemonic state during the instruction interval could be attributed to the small amount of training data. Second, we expect there to exist a single, ubiquitous retrieval state which, if preparatory in nature, should be engaged during both the instruction and stimulus interval. Such a view is in line with assumptions made in behavioral work demonstrating a stimulus-driven lingering influence of mnemonic states (Duncan et al., 2012; Patil & Duncan, 2018). According to this work, hippocampally mediated mnemonic states underlie the effect whereby responses on a prior trial influence memory accuracy on the current trial.

However, it is possible that distinct states are engaged during the instruction and stimulus intervals. Therefore, we conducted an exploratory analysis in which we performed leave-one-participant-out cross-validated classification on the instruction interval data. All data are derived from the current study. We performed classification across participants to overcome the limited number of trials available from individual participants; each classifier had access to 2052 (54 trials × 38 participants) training trials. We conducted classification on z-scored spectral power averaged across the instruction interval separately for each SOA condition. We averaged over time given that each SOA condition contributed a different number of time intervals.

We did not find significantly above-chance decoding for any SOA condition (500 ms:  $M = 50.28$ ,  $SD = 6.070$ ,  $t_{38} = 0.4261$ ,  $p = 0.6725$ ,  $d = 0.0946$ ; 1000 ms:  $M = 48.29$ ,  $SD = 5.862$ ,  $t_{38} = -1.865$ ,  $p = 0.0699$ ,  $d = 0.4259$ ; 1500 ms:  $M = 51.09$ ,  $SD = 6.088$ ,  $t_{38} = 1.143$ ,  $p = 0.26$ ,  $d = 0.2689$ ; 2000 ms:  $M = 48.81$ ,  $SD = 6.664$ ,  $t_{38} = -1.214$ ,  $p = 0.2324$ ,  $d = 0.2779$ ; Fig. 7). Thus, even when the classifier is trained on the instruction interval, spectral patterns do





**Fig. 7.** Leave-one-participant-out instruction-interval classification accuracy by SOA. We performed leave-one-participant-out within-study classification on the instruction interval, and derived classification accuracy for each SOA condition. The dotted horizontal line indicates permutation (chance-level) accuracy for each SOA condition. Error bars reflect standard error of the mean.

not consistently distinguish encode and retrieve trials. These findings provide further evidence against the hypothesis that the retrieval state is preparatory.

Together, our neural results support the reactionary account. We find robust evidence that participants selectively engage in memory encoding and memory retrieval after the stimulus onset. We do not find evidence that instruction influences retrieval state engagement during the instruction interval. We do not find evidence that SOA affects the impact of instruction on retrieval state engagement during the stimulus interval. These results show that participants engage mnemonic states following the stimulus onset, rather than in response to the instruction cue.

#### 4. DISCUSSION

The goal of this study was to investigate the time course of retrieval initiation. Using cross-study decoding of scalp EEG data (Long, 2023), we measured retrieval state engagement during a task in which participants were explicitly instructed to either encode the present event or retrieve a past, overlapping event. We find behavioral evidence that participants' ability to follow these instructions is unaffected by the stimulus onset asynchrony (SOA) between the instruction and stimulus. We find robust neural evidence that mnemonic brain states are selectively engaged after stimulus onset, and that this selectivity is not modulated by SOA. Together, these findings suggest that the retrieval state may constitute internal attention which requires a target, proposing a link between mnemonic brain states and attentional processes.

We find that the tendency to follow instructions to selectively encode vs. retrieve present vs. past stimuli is

unaffected by the latency between the instruction and stimulus. Consistent with our prior work (Long & Kuhl, 2019; Smith et al., 2022), we find higher memory performance for currently presented items that are presented with an “encode” instruction relative to items presented with a “retrieve” instruction. These findings are in line with the idea that encoding and retrieval trade off (Hasselman, 2005; Hasselman et al., 2002) such that engaging in retrieval precludes engaging in encoding (Long & Kuhl, 2019). To the extent that participants engage mnemonic states in response to the instruction cue (Evans et al., 2015; Morcom & Rugg, 2002; Rugg & Wilding, 2000), we would anticipate a larger dissociation in memory performance with longer SOAs. As longer preparatory intervals facilitate establishing task-appropriate cognitive states (Imburgio & Orr, 2021), longer instruction–stimulus intervals should provide additional time to recruit the goal-relevant state. However, we failed to find any effect of SOA, and even for the shortest SOA, we find a dissociation in memory performance on the basis of instructions. Although we cannot draw strong conclusions from behavioral data alone, the robust behavioral effects across SOA suggest that mnemonic states are not selectively recruited during the instruction interval.

The retrieval state is consistently, selectively engaged during the stimulus interval regardless of SOA. A preparatory retrieval state should be engaged prior to stimulus onset (Evans et al., 2015; Morcom & Rugg, 2002) such that we should find retrieval state dissociations across all time points in the stimulus interval and to a greater degree for longer SOAs. That is, the magnitude of the difference in retrieval evidence between encode and retrieve trials should increase as SOA increases. However, we find robust retrieval state engagement in response to the retrieve instruction after stimulus onset for all SOAs. Such findings are inconsistent with preparatory mnemonic states, but do follow from our prior work in which we demonstrated retrieval state dissociations around 1000 ms after stimulus onset (Long & Kuhl, 2019; Smith et al., 2022). Furthermore, these findings are in line with the interpretation that the retrieval state reflects internal attention (Chun et al., 2011; Logan et al., 2021; Long, 2023; Smith & Long, 2024). According to this framework, episodic retrieval can be thought of as the process of turning the mind's eye inward in the attempt to access a stored representation. A critical aspect of this framework is that a “target” is needed upon which attention can be focused—it is only possible to turn the mind's eye inward once a target, in our study the visually presented stimulus, is provided. Thus, retrieval state dissociations should only be observed during the stimulus interval.

We find that retrieval state evidence during the instruction interval is modulated only by time and not by instruction. To the extent that the retrieval state is preparatory,

we should have observed instruction-based dissociations whereby retrieval state evidence is greater for retrieve than for encode trials, and this effect should have been larger for longer SOAs. However, for all but one SOA, we failed to find an effect of instruction on retrieval evidence. In the remaining SOA, the effect is in the opposite direction of what we would have predicted—namely, greater retrieval state evidence on *encode* rather than *retrieve* trials. Thus, these findings do not support the preparatory account. We corroborated these findings by training and testing multivariate pattern classifiers on the instruction-interval data itself. The cross-study mnemonic state classifier that we used was, by design, trained on stimulus-interval data. Although we would expect a single retrieval state to be engaged during both instruction and stimulus intervals, it is possible that distinct states are recruited. However, we find no evidence that encode and retrieve instructions are dissociable during the instruction interval, even when a classifier is trained on instruction-interval data.

Our results appear to contradict past findings that demonstrate instruction-interval dissociations. Prior work on the retrieval state has demonstrated a positive voltage deflection over right prefrontal cortex approximately 600 ms following instruction onset (Morcom & Rugg, 2002). One potential avenue to reconcile these discrepancies is to consider the retrieval state in contrast to retrieval orientation. Whereas the retrieval state is characterized as a “tonic state that is maintained while episodic retrieval is required and remains constant across different episodic retrieval tasks” (Herron & Wilding, 2004), retrieval orientation should “vary according to the content of what episodic information is to be retrieved from memory” (Herron & Wilding, 2004). Extant studies have largely (though not exclusively, see Evans et al., 2015) compared episodic with semantic retrieval tasks based on the premise that the retrieval state is specific to episodic memory. Differences in the degree to which the retrieval state is preparatory may in part be due to the tasks being compared and how those tasks are operationalized. The preparatory effects previously observed may relate to participants’ ability to (at least partially) perform the task prior to probe onset. For example, when preparing to perform a size judgment task, participants could imagine a shoebox selectively for semantic vs. episodic trials. Thus, prior dissociations may be more closely related to the content of the upcoming task to be performed, rather than a content-general retrieval mode. Furthermore, given the suggestion that the retrieval state may reflect internal attention, it should be recruited in the service of any internal attention demands, whether episodic or semantic (Buckner & DiNicola, 2019). Therefore, one intriguing possibility is that the episodic/semantic

contrast is more akin to a superordinate retrieval orientation, rather than a retrieval state per se. This is an important direction for future work which we are actively pursuing. Finally, the ability to observe a tonically maintained, preparatory retrieval state may be dependent on the structure of the task. In contrast to the current study in which memory state instructions were randomly interleaved, prior work has found evidence for a preparatory retrieval state when comparing episodic and semantic retrieval tasks which were blocked into four-item lists (Duzel et al., 1999).

It is important to consider both that a preparatory retrieval state cannot fully account for evidence demonstrating automatic retrieval state engagement (Smith et al., 2022) and that evidence for the right frontal positivity is mixed. For example, as you prepare tonight’s dinner of boeuf bourguignon, you may find yourself reminded of a visit to France in 2013, without ever having had the intention of retrieving said trip. A stimulus-driven retrieval state can account for the finding that temporally overlapping experiences induce the retrieval state independently from top-down demands (Smith et al., 2022). Additional consideration of the robustness of the right frontal positivity is also warranted, given that reports of this event-related potential are somewhat inconsistent. The right frontal positivity has been selectively observed either only on “stay” trials, when the instruction (e.g. “retrieve”) is the same as the prior trial (Morcom & Rugg, 2002), or only on “switch” trials, when the instruction has changed from the prior trial (Evans et al., 2015). In theory, the retrieval state—insofar as it supports episodic remembering—should be recruited on both switch and stay trials. In the present study, we did not directly manipulate switch vs. stay trials and instead allowed instructions to randomly vary within a run such that participants will have encountered both types of trials. Even with this random variation, we find robust retrieval state engagement late in the stimulus interval.

In line with initial proposals of the retrieval state (e.g., Rugg & Wilding, 2000; Tulving, 1983), our interpretation is that the retrieval state is independent from retrieval success, based on both our current and prior findings. To reduce potential motor and/or decision-making related confounds, we do not require any behavioral responses during the encode/retrieve instructions. Thus, we do not directly measure retrieval success in the moment when participants are instructed to retrieve a previously presented stimulus. However, that we find dissociations in recognition accuracy based on the instruction to encode vs. retrieve suggests that participants attempt to engage the instructed mnemonic states. To the extent that memory encoding and memory retrieval tradeoff (Hasselman, 2005; Hasselman et al., 2002), engaging in retrieval—even

if unsuccessful—precludes engaging in encoding. A reduced ability or tendency to engage encoding will negatively impact subsequent memory for a currently presented, to-be-encoded stimulus. That these behavioral effects appear stronger for participants who show robust mnemonic state decoding suggests that engagement in mnemonic states impacts subsequent memory, consistent with our prior work (Long & Kuhl, 2019). Furthermore, we have also previously demonstrated that when participants are asked to provide behavioral responses in an explicit retrieval task, retrieval state evidence is greater for correct rejection of lures than for hits (Wheelock & Long, 2024) and for low compared with high confidence hits (Smith & Long, 2024). In each of these cases, there is more retrieval evidence when less information is available to support retrieval success. Together, these prior and present results suggest that the retrieval state reflects internal attention recruited to support the attempt to retrieve, regardless of whether said attempt is later successful.

Although we do not find an effect of instruction, we do find that SOA modulates instruction-interval retrieval state evidence. Specifically, we find less retrieval state evidence for shorter SOA conditions. We have intentionally built a classifier which places encoding and retrieval on a continuum based on the hypothesis that encoding and retrieval tradeoff (Hasselman, 2005; Hasselman et al., 2002; Long & Kuhl, 2019; Smith et al., 2022). Thus, given the design of our classifier, less retrieval state evidence corresponds to greater encoding state evidence. This in turn means that we find greater encoding state evidence for shorter SOAs during the instruction interval. There is an initial increase in encoding evidence in the 500 ms following cue onset across all SOAs, and for the shortest SOA, we find significant encoding state effects when we average over time. We additionally find robust decreases in retrieval state evidence early in the stimulus interval. Thus, following stimulus onset, participants initially engage the encoding state both when they are instructed to encode and when they are instructed to retrieve. By design, even for trials with a “retrieve” instruction, participants must perceive the current item to know which past item to attempt to retrieve. This perceptual processing may be reflected by the robust encoding state engagement immediately following stimulus onset. After participants initially engage the encoding state in response to cue and stimulus onset, encoding state engagement diminishes as participants continue to perceive the probe. Our interpretation is that encoding state evidence reflects *external* attention (Chun et al., 2011)—both to the instruction cue and to the stimulus, regardless of top-down goals.

In conclusion, our findings suggest that the retrieval state is reactionary and sustained over time. We find consistent, selective engagement of retrieval state evidence

on the basis of instructions during the stimulus interval, but not during the instruction interval, across all SOA conditions. These findings suggest that the attempt to access stored representations occurs after stimulus onset and is consistent with prior findings (Logan et al., 2021, 2024; Long, 2023) which suggest that the retrieval state may constitute internal attention. Together, these findings advance our understanding of the relationship between attention and memory, and how retrieval may be initiated.

## DATA AND CODE AVAILABILITY

The datasets generated in the current study, along with all experimental codes used for data collection and data analysis, are available on the Open Science Foundation (OSF; [osf.io/qmzfv](https://osf.io/qmzfv)) and via the Long Term Memory Lab website (<http://longtermmemorylab.com/papers/>).

## AUTHOR CONTRIBUTIONS

Subin Han: Data curation, Formal analysis, Software, Visualization, Writing—Original draft, Writing—Review & editing. Nicole M. Long: Conceptualization, Formal analysis, Funding acquisition, Project administration, Software, Supervision, Visualization, Writing—Original draft, Writing—Review & editing.

## FUNDING

This work was supported by a grant from the National Institutes of Health (NINDS R01 NS132872, PI: N.M.L.).

## DECLARATION OF COMPETING INTEREST

The authors declare no competing interests.

## ACKNOWLEDGMENTS

We thank Hannah Buras for assistance with data collection.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available with the online version here: <https://doi.org/10.1162/IMAG.a.1023>

## REFERENCES

- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, 11(5), 164–169. <https://doi.org/10.1111/1467-8721.00192>



- Buckner, R. L., & DiNicola, L. M. (2019). The brain's default network: Updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience*, 20, 593–608. <https://doi.org/10.1038/s41583-019-0212-7>
- Castellanos, N. P., & Makarov, V. A. (2006). Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods*, 158, 300–312. <https://doi.org/10.1016/j.jneumeth.2006.05.033>
- Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62, 73–101. <https://doi.org/10.1146/annurev.psych.093008.100427>
- DuBrow, S., & Davachi, L. (2014). Temporal memory is shaped by encoding stability and intervening item reactivation. *The Journal of Neuroscience*, 34(42), 13998–14005. <https://doi.org/10.1523/jneurosci.2535-14.2014>
- Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's penumbra: Episodic memory decisions induce lingering mnemonic biases. *Science*, 337(6093), 485–487. <https://doi.org/10.1126/science.1221936>
- Duzel, E., Cabeza, R., Ricton, T. W., Yonelinas, A. P., Scheich, H., Heinze, H.-J., & Tulving, E. (1999). Task-related and item-related brain processes of memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 1794–1799. <https://doi.org/10.1073/pnas.96.4.1794>
- Evans, L. H., Williams, A. N., & Wilding, E. L. (2015). Electrophysiological evidence for retrieval mode immediately after a task switch. *NeuroImage*, 108, 435–440. <https://doi.org/10.1016/j.neuroimage.2014.12.068>
- Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *The Journal of Neuroscience*, 30(29), 9793–9800. <https://doi.org/10.1523/jneurosci.0915-10.2010>
- Hasselmo, M. E. (2005). What is the function of hippocampal theta rhythm? - Linking behavioral data to phasic properties of field potential and unit recording data. *Hippocampus*, 15, 936–949. <https://doi.org/10.1002/hipo.20116>
- Hasselmo, M. E., Bodelon, C., & Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: Separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Computation*, 14, 793–817. <https://doi.org/10.1162/089976602317318965>
- Herron, J. E., & Wilding, E. L. (2004). An electrophysiological dissociation of retrieval mode and retrieval orientation. *NeuroImage*, 22, 1554–1562. <https://doi.org/10.1016/j.neuroimage.2004.04.011>
- Hong, Y., Moore, I. L., Smith, D. E., & Long, N. M. (2023). Spatiotemporal dynamics of memory encoding and memory retrieval states. *Journal of Cognitive Neuroscience*, 35(9), 1–15. [https://doi.org/10.1162/jocn\\_a\\_02022](https://doi.org/10.1162/jocn_a_02022)
- Imburgio, M. J., & Orr, J. M. (2021). Component processes underlying voluntary task selection: Separable contributions of task-set inertia and reconfiguration. *Cognition*, 212, 104865. <https://doi.org/10.1016/j.cognition.2021.104865>
- Konkle, T., Brady, T. F., Alvarez, G. A., & Aude, O. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *The Journal of Experimental Psychology: General* 139(3), 558–578. <https://doi.org/10.1037/a0019165>
- Kuhl, B. A., & Chun, M. M. (2014). Successful remembering elicits event-specific activity patterns in lateral parietal cortex. *The Journal of Neuroscience*, 34(23), 8051–8060. <https://doi.org/10.1523/jneurosci.4328-13.2014>
- Lewis-Peacock, J. A., & Norman, K. A. (2014). Competition between items in working memory leads to forgetting. *Nature Communications*, 5, 5768. <https://doi.org/10.1038/ncomms6768>
- Logan, G. D., Afu, K. C. S., Haynes, B. E., Weeks, E. E., Ulrich, J. E., & Lilburn, S. D. (2024). Attention focused on memory: The episodic flanker effect with letters, words, colors, and pictures. *Attention, Perception, and Psychophysics*, 86, 2690–2706. <https://doi.org/10.3758/s13414-024-02965-9>
- Logan, G. D., Cox, G. E., Annis, J., & Lindsey, D. R. B. (2021). The episodic flanker effect: Memory retrieval as attention turned inward. *Psychological Review*, 128(3), 397–445. <https://doi.org/10.1037/rev0000272>
- Long, N. M. (2023). The intersection of the retrieval state and internal attention. *Nature Communications*, 14(3861), 1–13. <https://doi.org/10.1038/s41467-023-39609-9>
- Long, N. M., & Kuhl, B. A. (2018). Bottom-up and top-down factors differentially influence stimulus representations across large-scale attentional networks. *The Journal of Neuroscience*, 38(10), 2495–2504. <https://doi.org/10.1523/jneurosci.2724-17.2018>
- Long, N. M., & Kuhl, B. A. (2019). Decoding the tradeoff between encoding and retrieval to predict memory for overlapping events. *NeuroImage*, 201, 1–10. <https://doi.org/10.1016/j.neuroimage.2019.07.014>
- Long, N. M., & Kuhl, B. A. (2021). Cortical representations of visual stimuli shift locations with changes in memory states. *Current Biology*, 31, 1119–1126. <https://doi.org/10.1016/j.cub.2021.01.004>
- Long, N. M., Kuhl, B. A., & Chun, M. M. (2018). Memory and attention. In J. T. Wixted (Ed.), *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*. (4th ed., Vol. 1, pp. 1–37). John Wiley & Sons Inc. <https://doi.org/10.1002/9781119170174.epcn109>
- Morcom, A. M., & Rugg, M. D. (2002). Getting ready to remember: The neural correlates of task set during recognition memory. *NeuroReport*, 13(1), 149–152. <https://doi.org/10.1097/00001756-200201210-00034>
- Nolan, H., Whelan, R., & Reilly, R. B. (2010). Faster: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, 192, 152–162. <https://doi.org/10.1016/j.jneumeth.2010.07.015>
- Patil, A., & Duncan, K. (2018). Lingering cognitive states shape fundamental mnemonic abilities. *Psychological Science*, 29(1), 45–55. <https://doi.org/10.1177/0956797617728592>
- Rugg, M. D., & Wilding, E. L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Sciences*, 4(3), 108–115. [https://doi.org/10.1016/s1364-6613\(00\)01445-5](https://doi.org/10.1016/s1364-6613(00)01445-5)
- Smith, D. E., & Long, N. M. (2024). Top-down task goals induce the retrieval state. *The Journal of Neuroscience*, 44(31), 1–11. <https://doi.org/10.1523/jneurosci.0452-24.2024>
- Smith, D. E., Moore, I. L., & Long, N. M. (2022). Temporal context modulates encoding and retrieval of overlapping events. *The Journal of Neuroscience*, 42(14), 3000–3010. <https://doi.org/10.1523/jneurosci.1091-21.2022>
- Tulving, E. (1983). *Elements of episodic memory*. Oxford University Press.
- Wheelock, J. R., & Long, N. M. (2024). Prior memory responses modulate behavior and brain state engagement. *Communications Psychology*, 2(1), 121. <https://doi.org/10.1038/s44271-024-00165-7>