Cognition 235 (2023) 105408

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Attentional fluctuations and the temporal organization of memory

Manasi Jayakumar *, Chinmayi Balusu, Mariam Aly

Department of Psychology, Columbia University, New York, NY 10027, United States of America

ARTICLE INFO

Keywords:
Long-term memory
Episodic memory
Temporal context
Event segmentation
Sustained attention

ABSTRACT

Event boundaries and temporal context shape the organization of episodic memories. We hypothesized that attentional fluctuations during encoding serve as "events" that affect temporal context representations and recall organization. Individuals encoded trial-unique objects during a modified sustained attention task. Memory was tested with free recall. Response time variability during the encoding tasks was used to characterize "in the zone" and "out of the zone" attentional states. We predicted that: 1) "in the zone", vs. "out of the zone", attentional states should be more conducive to maintaining temporal context representations that can cue temporally organized recall; and 2) temporally distant "in the zone" states may enable more recall "leaps" across intervening items. We replicated several important findings in the sustained attention and memory fields, including more online errors during "out of the zone" vs. "in the zone" attentional states and recall that was temporally structured. Yet, across four studies, we found no evidence for either of our main hypotheses. Recall was robustly temporally organized, and there was no difference in recall organization for items encoded "in the zone" vs. "out of the zone". We conclude that temporal context serves as a strong scaffold for episodic memory, one that can support organized recall even for items encoded during relatively poor attentional states. We also highlight the numerous challenges in striking a balance between sustained attention tasks (long blocks of a repetitive task) and memory recall tasks (short lists of unique items) and describe strategies for researchers interested in uniting these two fields.

1. Introduction

Episodic memories are temporally organized. Recall of a given event acts as a cue that can lead to recall of other events that were encoded close to it in time (Healey, Long, & Kahana, 2019; Howard & Kahana, 2002b; Kahana, 1996). This is thought to occur because events encoded close to one another in time have similar internal "temporal context" representations (Howard & Kahana, 2002b). Although temporal contexts powerfully shape recall, the factors that drive how those temporal contexts form during encoding are relatively underexplored. A better understanding of those mechanisms can help individuals promote factors that create strong temporal contexts at encoding and thus enhance memory retrieval. Here, we test the hypothesis that natural fluctuations in attention during encoding contribute to temporal context representations.

The Temporal Context Model (TCM) was proposed to explain the temporal organization of episodic memory (Howard & Kahana, 2002b). According to this model, at encoding, item representations are linked to a slowly changing but ever-present temporal context. When an item is recalled, the temporal context from encoding is also retrieved. Using the retrieved temporal context as a cue, other items studied with an overlapping temporal context can then be recalled (Howard & Kahana, 2002b; Polyn & Kahana, 2008).

Free recall studies examining these effects typically use lag-Conditional Response Probability (lag-CRP) curves to characterize the temporal organization of recall (Diamond & Levine, 2020; Healey et al., 2019; Howard & Kahana, 2002b; Kahana, 1996; Palombo, Di Lascio, Howard, & Verfaellie, 2019). These lag-CRP curves exemplify two characteristic features of temporally structured recall: the temporal contiguity effect and the forward asymmetry bias. The temporal contiguity effect refers to the tendency for items encoded close in time to be recalled close together (Healey et al., 2019; Howard & Kahana, 1999). The forward asymmetry bias refers to the higher probability of successively recalling items in a forward vs. backward direction. That is, for a given recalled item, subsequently recalled items are more likely to have been encoded after (rather than before) the first-recalled item (Polyn &

* Corresponding author at: Columbia University, Department of Psychology, 406 Schermerhorn Hall, 1190 Amsterdam Avenue MC 5501, New York, NY 10027, United States of America.

E-mail address: manasi.jayakumar@columbia.edu (M. Jayakumar).

https://doi.org/10.1016/j.cognition.2023.105408
Received 3 February 2022; Received in revised form 9 February 2023; Accepted 11 February 2023
0010-0277/© 2023 Elsevier B.V. All rights reserved.

Kahana, 2008; Polyn, Norman, & Kahana, 2009). This forward asymmetry bias is thought to arise because a given item becomes part of the temporal context for succeeding items, and can thus serve as a memory cue.

This research has largely proposed that temporal contexts are ever-present and slowly drift during an encoding experience, as newly encountered items and thoughts are incorporated into the temporal context representation. Whether certain cognitive and environmental factors can affect the drift of temporal context is an important question, because such factors may promote (or hinder) memory retrieval (DuBrow, Rouhani, Niv, & Norman, 2017). Recent research has shown that event segmentation can serve as one such factor that impacts the temporal organization of memory. Event segmentation theory hypothesizes that our ongoing experience is parsed into "segments" with the transition between segments acting as an event boundary (Clewett & Davachi, 2017; Heusser, Ezzyat, Shiff, & Davachi, 2018; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Memory retrieval is strongly shaped by such event boundaries: the order of events is remembered better, and events are more likely to be remembered as being closer together, if they were experienced within the same event segment vs. different event segments (DuBrow & Davachi, 2013, 2016; Heusser et al., 2018; also see Ezzyat & Davachi, 2010; DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014).

We propose that similar to event boundaries, natural fluctuations in attention may contribute to how temporal contexts form during encoding. Fluctuations in attention are inherent aspects of human nature (Killingsworth & Gilbert, 2010). People tend to experience times when attention peaks, leading to intense focus on the task at hand, while at other times, attention wanes and focus is broken by intrusive thoughts, distractions, or fatigue (Smallwood, Beach, Schooler, & Handy, 2008). Despite the ubiquity of such attentional fluctuations, past work examining how attention affects memory has typically focused on experimental manipulations of attention rather than spontaneous fluctuations. Such work has shown that experimental manipulations (such as directing participants' external attention towards a specific object, color, or spatial location) improve later memory for the attended event, but hurt memory for unattended events (Anderson, Craik, & Naveh-Benjamin, 1998; Craik, Govoni, Naveh-Benjamin, & Anderson, 1996, Craik, Eftekhari, & Binns, 2018; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Troyer, Winocur, Craik, & Moscovitch, 1999; Troyer & Craik, 2000; Yi & Chun, 2005; Chun & Turk-Browne, 2007; Chun, Golomb, & Turk-Browne, 2010; Uncapher, Hutchinson, & Wagner, 2011; Turk-Browne, Golomb, & Chun, 2013; LaRocque et al., 2015; see Aly & Turk-Browne, 2017).

A different body of work, on mind-wandering, has examined natural fluctuations of attention (Smallwood & Schooler, 2006; Christoff, Irving, Fox, Spreng, & Andrews-Hanna, 2016; also see Smallwood & Schooler, 2015). In these studies, participants are asked, with intermittent probes, to self-report whether they were "on task" or "mind-wandering" (e.g., Garlitch & Wahlheim, 2020; Metcalfe & Xu, 2016; Xu, Friedman, & Metcalfe, 2018; Xu & Metcalfe, 2016) or asked to describe their thoughts at the time of the probe (Smallwood, Baracaia, Lowe, & Obonsawin, 2003). These fluctuations in attention impact subsequent memory, such that mind wandering is related to worse memory (Garlitch & Wahlheim, 2020; Martarelli & Ovalle-Fresa, 2021; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Smallwood et al., 2003). However, these studies are limited in that they only capture participants' attentional state at a few discrete time points and cannot precisely characterize the temporal dynamics of intrinsic fluctuations in attention. Furthermore, such studies have not determined how mind wandering may affect the temporal structure of memory.

Sustained attention research offers a way to measure moment-by-moment fluctuations in attention (deBettencourt, Cohen, Lee, Norman, & Turk-Browne, 2015; deBettencourt, Norman, & Turk-Browne, 2018; Decker, Finn, & Duncan, 2020; Elshiekh & Rajah, 2021; Esterman, Noonan, Rosenberg, & Degutis, 2013; Esterman & Rothlein, 2019; Fortenbaugh, Rothlein, McGlinchey, DeGutis, & Esterman, 2018; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; Rosenberg, Finn, Scheinost, Constable, & Chun, 2017; Rosenberg, Noonan, DeGutis, & Esterman, 2013; Sarter, Givens, & Bruno, 2001; Smallwood & Schooler, 2006; Wakeland-Hart, Cao, deBettencourt, Bainbridge, & Rosenberg, 2022). One type of task, the gradual onset continuous performance task (gradCPT), uses reaction time (RT) variability to index moment-by-moment fluctuations: trials with higher RT variability constitute "out of the zone" attentional states, while trials with lower RT variability reflect "in the zone" attentional states. Online task performance differs based on these attentional states: participants make more errors in the task during an "out of the zone" attentional state (Esterman et al., 2013; Esterman, Rosenberg, & Noonan, 2014; Fortenbaugh et al., 2018; Rosenberg et al., 2013; Rosenberg et al., 2017). These studies, however, do not relate attentional fluctuations during the task to subsequent memory (see Madore et al., 2020 for a trait-level analysis). A related approach to characterizing attentional fluctuations based on RT showed that being in a good attentional state results in better recognition memory later on (deBettencourt et al., 2018; Wakeland-Hart et al., 2022), but these studies did not examine the temporal organization of memory. Building on these studies, in the current work, we use response time variability to characterize "in the zone" and "out of the zone" attentional states. We employed a modified version of the gradCPT, in which participants encoded trial-unique objects in the study phase (i.e., the sustained attention phase) and were later asked to verbally recall as many objects as they could in any order they chose. This allowed us to examine how moment-by-moment fluctuations in attention during encoding influence the temporal structure of recall.

We tested two main hypotheses. First, that recall will be more temporally structured when items are encoded "in the zone" vs. "out of the zone". This may occur because being focused on the task at hand leads to a consistent mental state, which can serve to bind and maintain a temporal context representation. Such focused attention may also facilitate linking attended events in the environment to these internal temporal contexts. Conversely, the reduction of task-focused thought in a "bad" attentional state may result in switches between internally and externally focused thoughts that disrupt a consistent temporal context representation, and/or hurt the ability to link items in the environment to internal temporal contexts. If this is the case, then the hallmarks of temporally organized memory — temporal contiguity and forward asymmetry — may be enhanced for items encoded "in the zone" vs "out of the zone".

Our second hypothesis is motivated by the finding that recall can "leap" between cognitively similar but temporally distant events (Chan, Applegate, Morton, Polyn, & Norman, 2017). This may occur because one's thoughts become integrated into temporal context representations: similar thoughts trigger similar temporal context representations, and in that way facilitate successive recall of items associated with cognitively similar contexts. We hypothesize that "in the zone" states, even if separated in time, constitute cognitively similar events: they consist of a focused mindset and the particular strategies that a person brings to mind to succeed in the ongoing task. On the other hand, we hypothesize that temporally distant "out of the zone" states are cognitively *dissimilar* events: every time an individual is unfocused, they may be unfocused in a different way, as their attention switches between the task and other ongoing, fluctuating thoughts. If this is the case, then recall may be more likely to "leap" between different "in the zone states" than different "out of the zone" states, bypassing items that were encoded in the other attentional state.

In sum, we aim to examine the behavioral effects of spontaneous attentional fluctuations on the temporal organization of recall. We characterize attentional fluctuations based on response time variability during an encoding task, classifying trials into relatively good "in the zone" and relatively worse "out of the zone" attentional states. To do that, we use a modified version of the gradCPT, introducing changes that make the task more suitable for examining subsequent memory. This

includes using trial-unique nameable objects, slower presentation durations, and fewer trials (see **Methods** - 2.1, 3.2, 4.2, 5.2 for more details). Participants' memory for the trial-unique objects was then tested with free recall. Lag-CRP curves, and analyses of recall based on different "in the zone" and "out of the zone" event segments, allowed us to test whether and how attentional fluctuations shape the temporal organization of memory.

It is worth emphasizing that addressing our question requires a synthesis of disparate paradigms that are used in the sustained attention and memory recall fields. To foreshadow our results, we found that striking the right balance between the opposing needs of these tasks can be a challenge — a challenge we attempted to tackle across 4 Studies. Our approach and results highlight the difficulties of studying the effects of attentional lapses on recall organization, and thus a secondary aim of our project is to raise awareness of these challenges and offer strategies for future research that tackles similar questions.

2. Study 1

2.1. Methods

2.1.1. Design

2.1.1.1. Participants. We conducted an a priori power analysis using G* Power (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). Due to the lack of prior work involving spontaneous attentional fluctuations and temporal organization of recall, we calculated the number of participants required to both a) replicate typical properties of the temporal organization of recall, including main effects of, and the interaction between, absolute lags and direction (Diamond & Levine, 2020; Kahana, Howard, Zaromb, & Wingfield, 2002; Palombo et al., 2019; Spillers & Unsworth, 2011); and b) observe interactions between lag-CRP properties (absolute lag or direction) and other independent variables (Palombo et al., 2019; Diamond & Levine, 2020; this was done to approximate interactions between lag-CRP properties and attentional state). For 80% power and an alpha of 0.05 in a within-participant design, we determined that the minimum required sample size was 50 participants. We opted to exceed that to counteract effect size overestimation resulting from publication bias (Bakker, van Dijk, & Wicherts, 2012; Brand, Bradley, Best, & Stoica, 2008). We therefore report data from 65 participants ($M_{age} = 25.17 \pm 6.67$ years, $M_{education} = 14.78 \pm 2.23$ years; see Table 1 for demographics). We do not report data from an additional 15 participants, who were excluded due to image loading errors ($N = 6$), low response rate during the encoding task (<80%, $N = 2$), outlier response accuracy during the encoding task (>3 SD from the group mean; $N = 3$), recall recording issues (N = 3), and incomplete data due to technical problems ($N = 1$).

Of the final sample, 22 participants were recruited from the Columbia University participant pool. They completed the study in the lab and were compensated with course credit. Because of the COVID-19 pandemic and the related disruption of research activities, the remaining participants ($N = 43$) were recruited through Prolific (www.prolific.co). They participated in an online version of the same experiment hosted on the Gorilla platform (www.gorilla.sc; Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020).

Participants were 18 to 40 years of age, fluent in English, and resided in the US (inclusion criteria were specified in Prolific prior to recruitment). Both groups of participants provided informed consent in accordance with the Columbia University Institutional Review Board. No statistically significant differences were observed between the in-person and Prolific samples in any measure of interest (all $ps > 0.14$ for all main effects and interactions involving the 'sample' variable); thus, all data analyses include the combined sample. Nevertheless, for completeness, we report statistics to compare the two groups for effects of interest.

Table 1
Summary of participant demographics across studies. For age and education, we report the mean ± standard deviation. For recruitment method, CU = Columbia University participant pool. For gender, F = female, M = male, NB = non-binary, DNS = did not specify. For race, W = white, A = Asian, B/AA = Black or African American, BR = bi-racial, AI/AN = American Indian or Alaskan Native, ME = Middle Eastern, O = other. For ethnicity, NH/L = not Hispanic or Latino, and H/L = Hispanic or Latino. In Study 4, three participants wrote in Latino for race.

	Study 1	Study 2	Study 3	Study 4
Sample size	65	68	68	124
Age (M ± SD)	25.17 ± 6.67	22.62 ± 5.15	20.09 ± 2.20	21.42 ± 6.26
Education (M ± SD)	14.78 ± 2.23	13.99 ± 1.65	14.03 ± 1.46	13.63 ± 1.54
Recruitment Method	22 CU 43 Prolific	48 CU 20 Prolific	68 CU	119 CU 5 Prolific
Gender	33 F 28 M 3 NB 1 DNS	44 F 23 M 1 NB	33 F 35 M	84 F 37 M, 2 NB 1 DNS
Race	33 W 16 A 8 B/AA 5 BR 1 AI/AN 1 ME 1 O	39 W 15 A 10 B/AA 4 BR	35 W 19 A 7 B/AA 4 BR 1 AI/AN 2 O	67 W 30 A 14 B/AA 4 BR 3 AI/AN 3 H/L 1 ME 2 DNS
Ethnicity	59 NH/L 6 H/L	60 NH/L 8 H/L	54 NH/L 14 H/L	98 NH/L 26 H/L

2.1.1.2. Stimuli. We chose 191 images of objects from pre-curated object databases such as SOLID (Frank, Gray, & Montaldi, 2020), stimuli from the Mnemonic Similarity Task (Yassa et al., 2011; https://faculty.sites.uci.edu/starklab/mnemonic-similarity-task-mst/), Interaction Envelope (Bainbridge & Oliva, 2015a, 2015b; http://www.wilmabainbridge.com/datasets.html), and the Bank of Standardized Stimuli (Brodeur, Guérard, & Bouras, 2014; https://sites.google.com/site/bosstimuli/). Color images were converted to grayscale.

These images were assigned to 5 study blocks using the OptSeg tool (Siegelman, 2019; https://github.com/msieg/OptSeg_Reproducible), which pseudo-randomizes the stimuli into lists while controlling for semantic similarity between constituent words. Semantic similarity between two items was measured as the cosine distance between the 300-dimensional GloVe vectors of the object names (Siegelman, 2019). Given a pool of words, the algorithm constructs lists so that items within a list are as semantically dissimilar as they can be. This allowed us to focus on temporal organization within lists that minimized opportunities for semantic organization given the stimuli available (Manning & Kahana, 2012). All 191 images were provided to the algorithm to have some leeway in assigning semantically matched and optimized lists. We created 5 lists of 30 images each using this procedure. The 5 stimulus lists were then randomized to 5 study blocks with within-block randomization of image order. The remaining 41 images were used for the practice block before the main task.

Although this approach attempts to minimize semantic similarity within a list, each list still had categorical structure. Specifically, within a list, items typically belonged to one of these categories: Musical Instruments, Animals, Plants, Food and Drinks, Tools, Home and Furniture, or Sports and Games. Thus, we could not rule out semantic clustering in our task, a topic we return to in the Section 6 - **General Discussion**.

2.1.1.3. Procedure. The experiment was composed of 5 blocks, each of which included a study phase, a distractor phase, and a recall phase (Fig. 1).

In each of the 5 study phases, participants viewed a series of 30 trial-unique images of common objects. The objects transitioned from one into another over 5 s. Each object remained on the screen for 1 s before

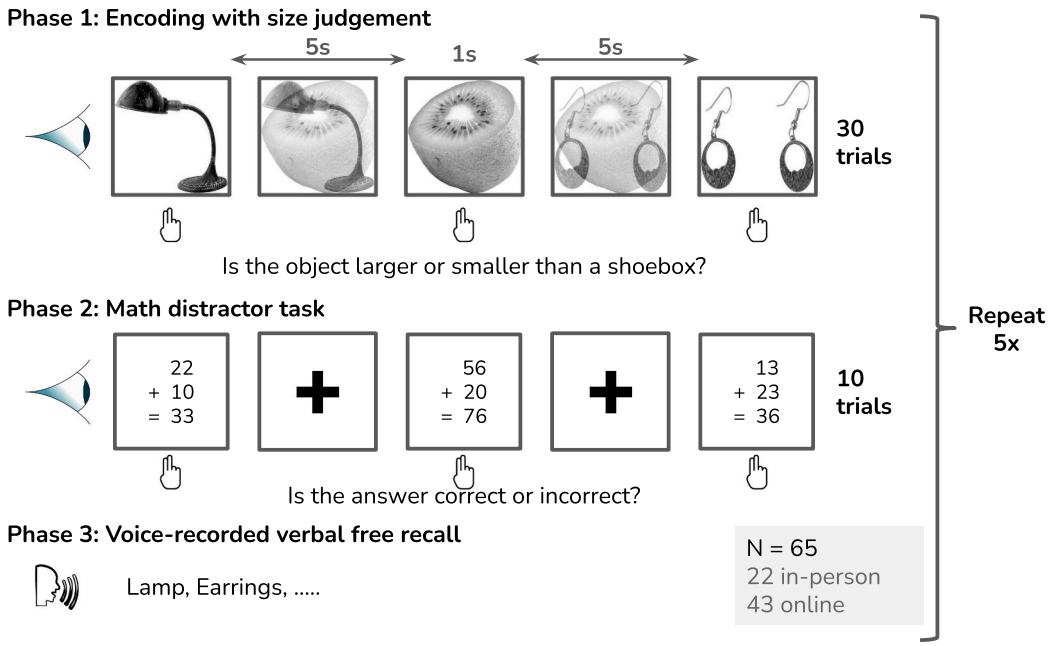

Fig. 1. Schematic of the experimental design for Study 1. Participants were presented with a series of grayscale objects in the study phase. They were asked to judge if each object is smaller or larger than a shoebox. Object images gradually transitioned from one into another. Following a math distractor task, participants were asked to verbally recall objects from the study phase, in any order that they choose. This 3-phase sequence was repeated in 5 blocks, with 30 images encoded each time. Study 2 was similar except that there were 3 blocks of 80 items each and the task was to press a button for non-food items and withhold a response for food items. Study 3 had 3 blocks of 60 items each and the task was to press a button for color images and withhold a response for grayscale images. In Study 3, there were no gradual transitions between items; each image was shown for 3 s with a 2 s inter-item interval. Study 4 was similar to Study 2, except that there were 2 blocks of 120 images each, and the trial duration was reduced from 6 s to 4 s.

the start of the next transition. Every 0.5 s the first object's opacity decreased by 10% while the succeeding object's opacity increased by 10%. This slow transition ensures that there is no capture of attention by abrupt image onsets, and helps induce more fluctuations of attention (Esterman et al., 2013; Rosenberg et al., 2013). For each presented image, participants were asked to judge if the depicted object was "smaller or larger than a shoebox" using one of two keys. These two keys were counterbalanced across participants. Participants could respond at any time once the object had started fading in (Fig. 1).

A distractor phase followed each study phase, to introduce a delay between memory encoding and subsequent recall. Participants were presented with a sequence of 2-digit additions and subtractions. One solution to each question was provided and participants had to judge if the solution was correct or incorrect using one of two keys. Each trial began with a fixation cross presented for 2 s, followed by the arithmetic question for 5 s. 10 arithmetic questions were presented in each phase.

Memory recall immediately followed this distractor task. During each recall phase, participants were asked to verbally recall, in any order, as many objects as they could from the preceding study phase. This phase was recorded using a microphone. There was a time limit of 2–2.5 min (see below), and a countdown timer was displayed to indicate the time remaining.

Before beginning the experimental blocks, participants performed a practice block. The practice block was identical to the experimental blocks except that only 15 objects were presented in the study phase.

Both versions of the task (in-person and online) were similar except for two minor modifications. The verbal recording for the in-person version was 2.5 min long for each recall portion; for the online version, it was 2 min long (the maximum allowed on Gorilla). Because participants rarely used the entire 2.5 min for recall in the in-person version, this was a minor change that had no measured impact on performance (overall recall was not significantly different between the in-person and online groups ($t_{48.18} = 1.32$, $p = 0.19$, Cohen's ds = 0.35, 95% CI [−2.53, 12.16]) nor was recall organization different between the groups, all $ps > 0.14$ for interactions involving sample [online vs. in-person] in lag-CRP and event segment analyses discussed below). Additionally, the practice items for the in-person version were chosen from the leftover images after the Optseg assignment, and different participants were shown a different set of images. However, for the online version all participants saw the same 15 images in a randomized order.

We used 30 items per block — more than many standard list-learning recall studies (see Sederberg, Miller, Howard, & Kahana, 2010 for a few examples) because we expected that longer blocks (i.e., longer lists of items) would be necessary to reliably induce attentional fluctuations. Indeed, much longer block lengths (400 trials or more) are used in studies of sustained attention (Decker et al., 2020; Esterman et al., 2013; Rosenberg et al., 2013; Wakeland-Hart et al., 2022). We, therefore, had fewer lists than typical list-learning recall studies, because more items were encoded per list. In Studies 2–4, we further varied the number of lists and the number of items within them, in the hope of striking a good balance between obtaining good recall (which is typically better with shorter lists) and inducing attentional fluctuations (which typically occurs with long blocks).

2.1.2. Analyses

2.1.2.1. Defining attentional states at encoding. Attentional states were defined using response times (RTs) for the judgements made during the study phase. To assign responses to trials, we used a similar procedure as that used by Esterman et al. (2013) and Rosenberg et al. (2013). One modification was made to accommodate the slower image transitions in our tasks (3 s or 5 s in our Studies rather than 800 ms in prior work). Furthermore, in our Studies, images stayed 100% coherent for 1 s (rather than no "pause" at 100% coherence or a 400 ms "pause" at 100% coherence in prior work). Because of these changes, there was a large peak in the RT distributions when an image was 90–100% coherent, and

the majority of responses (~83–93%, across our Studies) came when an image was fading in between 50 and 100% coherence. There were few responses when an image was fading out from 100% coherence because of the relatively long 1 s "pause" at the 100% coherent image (during which many responses were made). We therefore adjusted the image coherence thresholds used in prior work to reflect our observed RT distributions.

RTs were therefore assigned as follows. RTs were calculated relative to the beginning of each image transition (i.e., when it was starting to fade in). Responses when an image was fading in between 50% and 100% coherence were assigned to that image; this reflected the majority of trials. On the rarer trials for which a response occurred when an image was fading in between 0 and 50% (i.e., the previous image was fading out from <100% to 50%), an algorithm was applied as follows (following Esterman et al., 2013 and Rosenberg et al., 2013). First, the algorithm assigned unambiguous correct responses — that is, responses to the "go" category that occurred between 50% and 100% coherence of the image that was fading in. Second, the remaining, ambiguous responses (i.e., those from 0 to 50% coherence of the image fading in or multiple responses) were assigned to an adjacent trial if one of the two had no response. If both adjacent trials had no response, ambiguous responses were assigned to the closest trial, unless one was a "no-go" trial, in which case participants were given the benefit of the doubt that they had correctly omitted. Finally, if there were multiple responses that could be assigned on any one trial, the fastest response was selected. The variance time course analyses described below were also run with an alternative approach that simply assigned a response to the image that was currently fading in; the pattern of results was largely unaffected with this alternative procedure.

A variance time course (VTC) analysis was performed on the RT data using the procedure in Esterman et al., 2013. This procedure enables the identification of two attentional states ("in the zone" and "out of the zone") based on the variability of RT. Each correct trial was assigned a value corresponding to the absolute deviation of the trial RT from the mean within-block RT. Trials without a response and trials with an incorrect response were not included in this step. Next, this value was linearly interpolated (from the neighboring two trials) for trials without a response and trials with an incorrect response. If only one trial was available for interpolation (i.e., because trials at the beginning or end of a block do not have two surrounding trials), then RT was not interpolated and such trials were not assigned an attentional state. Then, following the procedure used by Esterman et al. (2013) and Rosenberg et al. (2013), a Gaussian smoothing kernel was applied. We used the *smth* function in the smoother R package (https://CRAN.R-project.org/package=smoother), and set the window to 4 trials (full-width at half-maximum) and method to Gaussian. This procedure therefore integrated information from the 4 surrounding trials (approximately 24 s, as in Rosenberg et al., 2013). Finally, a median split was performed on these smoothed variance time course values, dividing the trials into those with lower RT variability (i.e., RTs closer to the mean; "in the zone" states) and higher RT variability (i.e., RTs farther away from the mean; "out of the zone" states).

This VTC analysis therefore assumes that RTs that are too fast or too slow (with respect to their distance from the mean RT) reflect a poor attentional state. The rationale for this is that RTs that are too fast reflect habitual responding or being on "autopilot" and RTs that are too slow reflect disengagement from, or inefficient processing of, the task at hand (Esterman et al., 2013; Kane & McVay, 2012; Yamashita et al., 2021). Indeed, this approach has been successful at predicting attentional lapses (e.g. Esterman et al., 2013, 2014; Madore et al., 2020; Rosenberg et al., 2013, Rosenberg, Finn, Constable, & Chun, 2015; also see Yamashita et al., 2021). However, it is critical to note that there are other approaches to characterizing attentional states based on RT. For example, some studies of sustained attention have shown that faster (vs. slower) RTs are linked to more online errors and worse subsequent recognition memory (e.g., Cheyne, Carriere, & Smilek, 2006; Cheyne, Solman, Carriere, & Smilek, 2009; deBettencourt et al., 2018; Kane & McVay, 2012; Robertson et al., 1997; Wakeland-Hart et al., 2022). Yet, counter to these approaches, other studies link longer eye fixations and slower response times to mind wandering and attentional lapses (e.g., Smallwood et al., 2003; Weissman, Roberts, Visscher, & Woldorff, 2006; Feng, D'Mello, & Graesser, 2013; Yanko & Spalek, 2013; Kam & Handy, 2014; Henríquez, Chica, Billeke, & Bartolomeo, 2016; Krasich et al., 2018; Zhang, Anderson, & Miller, 2021). Thus, slower (vs. faster) RTs have sometimes been linked to better attentional states and sometimes to worse attentional states, and the VTC analysis considers RTs that are either too fast or too slow (relative to the mean) to reflect a bad attentional state.

To accommodate this diversity in identifying better vs. worse attentional states from RTs, we provide a secondary analysis for each study in which we examined attentional states defined by faster RTs vs. slower RTs. Although the VTC approach was our a priori analysis of choice, this allowed us to examine whether alternative definitions of attentional states lead to different results with respect to online errors and subsequent memory.

2.1.2.2. Errors at encoding. In prior studies, "out of the zone" (vs. "in the zone") attentional states were associated with more errors in sustained attention tasks (Esterman et al., 2013; Rosenberg et al., 2013). To determine if we could replicate those findings, we examined whether there was a difference between the two attentional states in the number of errors made during the encoding task, which was designed to be similar to the sustained attention tasks used in prior studies. First, objects depicted by each image were classified as being either larger than a shoebox (e.g., helicopter, treadmill), smaller than a shoebox (e.g., onion, key), or ambiguous (e.g., cowboy hat, soda bottle). Next, errors were calculated as the number of incorrect responses made to the objects that were unambiguously classified (i.e., responses to ambiguous objects were never counted as incorrect; Note that accuracy was therefore the percentage of presented items with a correct response, for which responses to ambiguous objects were always counted as correct). Finally, we examined whether the number of errors made during "out of the zone" attentional states was higher than the number of errors made during "in the zone" states. Group-level analyses were conducted with a paired-samples *t*-test.

2.1.2.3. Recall performance. We calculated recall performance as the percentage of items that were correctly recalled across all blocks. To examine whether recall performance differed by attentional state, we calculated recall performance for each attentional state ("in the zone" or "out of the zone") as the percentage of items that were encoded in that particular attentional state that were correctly recalled. Note that because these attentional states are defined by a median split of the encoding RTs, the same number of items are encoded in each state. Group-level analyses were conducted with a paired-samples t-test.

2.1.2.4. Temporal organization of recall. We measured the temporal organization of recall using lag-Conditional Response Probability (lag-CRP) curves (Kahana, 1996). The lag-CRP curve measures the probability of recalling two items successively as a function of their relative position, or lag, at encoding. To plot these curves, we first obtain the encoding lag between all pairs of successively recalled items, where the lag is the difference between their serial positions at encoding. The lag can be positive or negative, depending on whether the subsequent item recalled was encoded after (positive lag) or before (negative lag) the preceding item. The observed number of recall transitions at each lag is then divided by the number of opportunities to make a recall transition at that lag, e.g., all the times a participant could have recalled an item at a + 1 lag, regardless of whether or not they did (Kahana, 1996). This yields the probabilities plotted in the lag-CRP curve, i.e., actual transitions divided by possible transitions at each lag. Repetitions and

intrusions are masked from this analysis: transitions to and from repetitions (recalled items that had also been recalled earlier) and intrusions (items recalled from a prior study list) are excluded (Kahana, 1996).

Lag-CRP curves depict two characteristic features of the temporal organization of recall: forward asymmetry and temporal contiguity. Forward asymmetry refers to the higher likelihood of recalling in the forward vs. backward direction (i.e., higher conditional probability of recall for positive lags vs. negative lags). Temporal contiguity refers to the higher probability of recalling items together if they were encoded closer together in time (Healey et al., 2019; Kahana, 1996). This is seen as a peak in the lag-CRP curves: recall is more likely for items at ±1 lag, and recall probability decreases gradually with increasing lags.

Our primary hypothesis was that "in the zone", vs. "out of the zone", attentional states are more conducive to maintaining a temporal context representation. If so, there should be a difference in the temporal organization of recall between the attentional states: temporal contiguity and/or forward asymmetry should be stronger for "in the zone" vs. "out of the zone" attentional states. To test this, we constructed separate lag-CRP curves for the two attentional states. First, individual pairs of successively recalled items were labelled as being in the same state (i.e., both encoded during an "in the zone" state or both encoded during an "out of the zone" state) or being a transition pair from one state to another. To calculate the lag-CRP curves separately for each state, we only considered pairs that were in the same state; transition pairs were excluded from analysis (see Section 2.1.2.5: **Recall Transitions by Event Segment** for consideration of state transitions). Actual transitions were calculated as the lag between two successively recalled items based on their encoding position. These actual transitions were divided by the number of times a transition of a given lag could have possibly occurred irrespective of attentional states (i.e., all possible transitions) to yield the conditional response probability. The conditional response probability for each lag was calculated within-block for each attentional state. These values were then averaged across blocks, resulting in one CRP value at each lag for each attentional state for each participant.

To test for a statistical difference in the temporal organization of recall between the two attentional states, we performed a three-way repeated-measures ANOVA with attentional state ("in the zone" vs. "out of the zone"), absolute lag (1 to 29), and recall direction (forward vs. backward) as factors. We tested the sphericity assumption using Mauchly's test of sphericity and applied the Greenhouse-Geisser correction when this assumption was not satisfied. This analysis also allowed us to test for typical properties of lag-CRP curves, by looking for main effects of absolute lag and direction, and the interaction between them.

Separate lag-CRP curves for each attentional state (and the associated 3-way ANOVA) required us to discard some data (recall transitions *between* "in the zone" and "out of the zone" states and participants who did not have any successive recall transitions within "in the zone" states or within "out of the zone" states). Because of this, we also show the overall lag-CRP curves for each Study (using all the data, irrespective of attentional state at encoding), and report analyses of the overall lag-CRP curve with a two-way (direction x lag) ANOVA in Table 2.

Finally, because there is a possibility that including all lags could mask differences in forward asymmetry — which is typically most prominent at nearby lags — we conducted an additional analysis. We examined whether there was a difference between the two attentional states in forward asymmetry at the closest lags of ±1, by performing a two-way repeated-measures ANOVA with lag (+1 vs. -1) and attentional state ("in the zone" vs. "out of the zone") as factors (Diamond & Levine, 2020).

The approach we take in the above analysis — dividing actual transitions within an attentional state by all possible transitions at a given lag *regardless of attentional state at encoding* — effectively asks: of all possible recall transitions available to the participant at lag of X, how many were for items encoded "in the zone" vs. "out of the zone"? For example, we divide the number of +1 transitions between "in the zone" items by the total number of possible +1 transitions regardless of attentional state at encoding; because the denominator is therefore the same for the two attentional states, we will refer to this as the "same denominator" approach.

An alternative analysis choice would be to divide actual transitions by the number of times a transition of a given lag could have possibly occurred *within that same attentional state*; e.g., divide +1 transitions for "in the zone" items by the number of times a + 1 transition *to other "in*

Table 2
Summary of two-way ANOVA (direction x absolute lag) results for the overall lag-CRP curves in each Study. Green shading indicates statistically significant results ($p < 0.05$) and gray shading indicates results that are not statistically significant.

	Study 1	Study 2	Study 3	Study 4
Lag (1 to 29)	Nearby > far away $F_{12.42, 794.89} = 11.28$ $p < 0.0001$ $\eta_p^2 = 0.15$	Nearby > far away $F_{11.36, 761.14} = 14.47$ $p < 0.0001$ $\eta_p^2 = 0.18$	Nearby > far away $F_{11.53,772.73} = 11.49$ $p < 0.0001$ $\eta_p^2 = 0.15$	Nearby > far away $F_{13.46, 1655.35} = 7.34$ $p < 0.0001$ $\eta_p^2 = 0.06$
Direction (Forward vs. Backward)	$F_{1,64} = 0.33$ $p = 0.57$ $\eta_p^2 = 0.005$	Forward > backward $F_{1,67} = 15.77$ $p = 0.0002$ $\eta_p^2 = 0.19$	$F_{1,67} = 0.002$ $p = 0.97$ $\eta_p^2 = 0.00003$	$F_{1,123} = 1.35$ $p = 0.25$ $\eta_p^2 = 0.01$
Lag x Direction	$F_{14.79, 946.73} = 1.45$ $p = 0.12$ $\eta_p^2 = 0.02$	Forward > backward for nearby > far away $F_{15.85,1062.11} = 1.88$ $p = 0.019$ $\eta_p^2 = 0.027$	Forward > backward for nearby > far away $F_{15.86,1062.47} = 1.84$ $p = 0.023$ $\eta_p^2 = 0.027$	$F_{16.06,1975.52} = 1.27$ $p = 0.21$ $\eta_p^2 = 0.01$
+1 vs. -1 transitions	$t_{64} = 1.89$ $p = 0.063$ Cohen's dz = 0.23 95% CI [-0.002,0.059]	$t_{67} = 3.59$ $p = 0.0006$ Cohen's dz = 0.44 95% CI [0.01,0.04]	$t_{67} = 3.32$ $p = 0.002$ Cohen's dz = 0.40 95% CI [0.01,0.04]	$t_{123} = 1.92$ $p = 0.06$ Cohen's dz = 0.17 95% CI [-0.0003,0.018]

the zone" items could have occurred. We will refer to this as the "state-specific denominator" approach.

In simulated data, we found that these approaches sometimes yield identical *differences* between attentional states, although the *absolute value* of the CRP will necessarily be higher when the denominator is specific to a given attentional state. In other cases, the "state-specific denominator" approach underestimated differences in temporal structure between conditions, while in yet other cases, it was more accurate than the "same denominator" approach — particularly when recall rates differed between conditions but temporal structure did not. Because this "state-specific denominator" approach can be argued to be valid in some situations, we repeated all the analyses (across all Studies) with this alternative approach — but the pattern of results did not change (**Table S1**).

2.1.2.5. Recall transitions by event segment. The above temporal organization analyses do not differentiate between qualitatively different types of transitions that are possible within each attentional state. Because individuals fluctuate between "in the zone" and "out of the zone" periods during encoding, each instance of an attentional state can be considered its own "event segment" (Fig. 2). These different segments are ignored in the lag-CRP analysis above, which simply considers each attentional state as a whole. Furthermore, our "in the zone" and "out of the zone" lag-CRPs ignored recall transitions *between* "in the zone" and "out of the zone" states. Yet, considering these types of transitions is important: research on event boundaries has shown that successive recall of adjacently encoded items is more likely when those items are encoded in the same vs. different event (Ezzyat & Davachi, 2010; DuBrow & Davachi, 2013, 2016; Heusser et al., 2018; also see DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014) and recall may "leap" between cognitively similar but temporally distant events (Chan et al., 2017). We therefore designed an analysis to test whether fluctuations between attentional states can act as event boundaries that reproduce the phenomena observed in research on events.

This analysis allowed us to test whether 1) recall transitions within an "event segment" are more likely for "in the zone" vs. "out of the zone" attentional states; and 2) whether recall for "in the zone" states is more likely to "leap" from one event segment to another, bypassing items that were encoded in an "out of the zone" attentional state (Chan et al., 2017; also see Heusser et al., 2018). To examine these hypotheses, we considered all pairs of successively recalled items, and calculated the number of transitions made during recall for each type of transition noted in Fig. 2 (same state, same event; same state, different event; different state, different event), separately for each attentional state ("in the zone" vs. "out of the zone"). We divided the number of actual transitions in each bin by the number of opportunities to transition to another item that falls within the same transition type regardless of attentional state at encoding. The resulting value is therefore the conditional probability of each transition type. For example, we divided the number of "same state, same event" transitions for "in the zone" items by the number of opportunities to make "same state, same event" transitions to other items, regardless of attentional state at encoding (i.e., the total number of possible "same state, same event" transitions between two "in the zone" items and between two "out of the zone" items). This analysis therefore asks: of all possible event transitions of a specific type (e.g., "same state, same event"), how many were for items encoded "in the zone" vs. "out of the zone"? This is analogous to the "same denominator" approach we used for the lag-CRPs.

For completeness, we also re-ran this analysis using the "state-specific" denominator approach, in which we divided the number of transitions in each bin by the number of opportunities to transition to another item that falls within the same type of transition *for that attentional state*. For example, we divided the number of "same state, same event" transitions for "in the zone" items by the number of opportunities to make "same state, same event" transitions to items encoded "in the zone".

In both the "same denominator" and "state-specific denominator" approaches, these conditional probabilities were calculated separately for each block, and then averaged across blocks for each participant. We then performed a two-way repeated-measures ANOVA with the type of transition (3 levels) and attentional state (2 levels) as factors. The pattern of results was the same for the "same denominator" and "state-specific" denominator approaches; here, we report the results for the "same denominator" approach to be consistent with the lag-CRP analysis. The results for the "state-specific" denominator approach can be found in **Table S1**.

2.1.2.6. Bayes Factors. We computed Bayes Factors using the BayesFactor package in R (https://cran.r-project.org/web/packages/BayesFactor/index.html;Morey & Rouder, 2011; Rouder & Morey, 2011; Rouder, Morey, Speckman, & Province, 2012, Rouder, Morey, & Province, 2013). Bayes Factors quantify the strength of evidence provided by the data for one hypothesis vs. another.

We report Bayes Factors (BF) in terms of support for the null hypothesis, such that values >3 indicate evidence for the null while values <1/3 indicate evidence for the alternative hypothesis (Jeffreys, 1961;

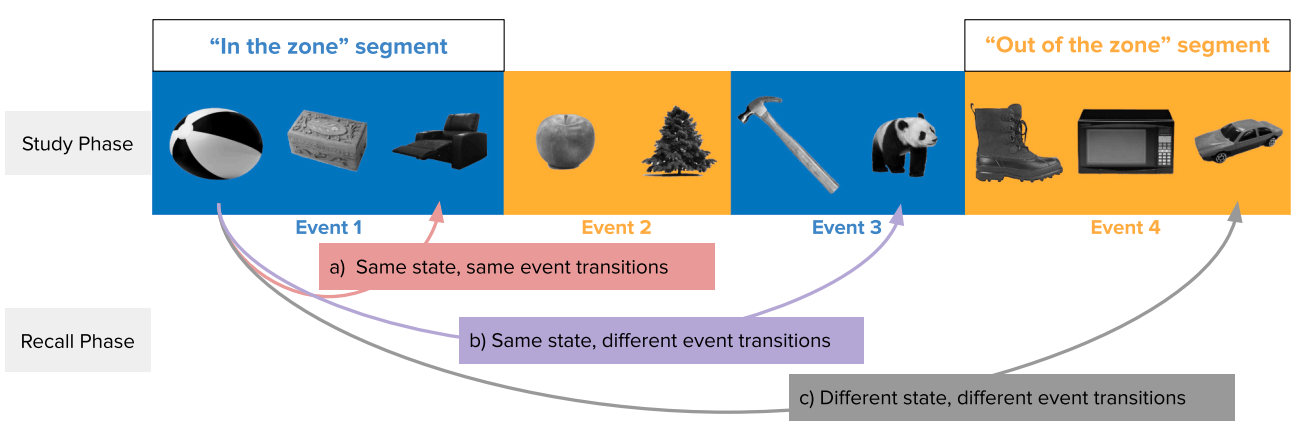

Fig. 2. Types of transitions at recall. Individuals fluctuate between "in the zone" (blue) and "out of the zone" (orange) attentional states during encoding. Each instance of each attentional state can be considered its own "event segment". For example, the ball and the recliner are encoded in the same attentional state and the same event segment within that state (**a**: "same state, same event"), whereas the ball and the panda are encoded in the same attentional state but during different event segments (**b:** "same state, different event"). The ball and the car are encoded in different attentional states and therefore, by necessity, different event segments as well (**c:** "different state, different event"). In the recall phase, given the recall of one object, the transition to the second object can be therefore categorized as one of three types: 1) same state, same event; 2) same state, different event; and 3) different state, different event. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Kass & Raftery, 1995; Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011). Specifically, Bayes Factors between 1/3–3 indicate no evidence for either the null or alternative hypothesis; Bayes Factors between 3 and 20 or between 1/20–1/3 provide evidence for the null or alternative hypothesis, respectively; Bayes Factors between 20 and 150 or between 1/150–1/20 provide strong evidence for the null or alternative hypothesis, respectively; and Bayes Factors >150 or <1/150 indicate very strong evidence in favor of the null or alternative hypothesis, respectively. For the sake of completeness and transparency, we report Bayes Factors for all of our main findings (in all studies and in Table 3). However, we note that care must be taken when computing and interpreting Bayes Factors because they can be subject to many limitations (e.g., sensitivity to prior selection), including limitations similar to those of *p*-values (e.g., binary treatment of a continuous measure). For discussions about the limitations of Bayes Factors, see Liu & Aitkin, 2008; Simonsohn, 2014; Simmons, Nelson, & Simonsohn, 2016; Gelman & Carlin, 2017.

Bayes Factors for paired *t*-tests were computed with the function *ttestBF()* with *paired* set to *TRUE*, and other arguments left as the default; the inverse was then obtained to provide the Bayes Factor in terms of support for the null. For main effects and interactions from ANOVAs, we computed Bayes Factors using the function *lmBF()* with default priors and 500,000 iterations. Bayes Factors for main effects were computed by dividing the Bayes Factor for the model excluding the main effect (and all interactions with the effect of interest) by the Bayes Factor for the model including the main effect, but no interactions with the effect of interest. Similarly, to compute the Bayes Factor for an interaction effect, we divided the Bayes Factor obtained from a model excluding the interaction effect (but including all other main effects and interactions) by the Bayes Factor for a model including the interaction of interest

Table 3
Summary of results across Studies 1–4. Green shading indicates statistically significant results ($p < 0.05$) and gray shading indicates results that are not statistically significant. The interpretation of Bayes Factors (BF) is described in the main text. F-values, t-values, degrees of freedom, effect sizes, and confidence intervals are reported in the main text. In = "in the zone" attentional state; Out = "out of the zone" attentional state, defined with the VTC analysis.

		Study 1	Study 2	Study 3	Study 4
	Online errors (In vs. Out)	Out > In p < 0.0001 BF < 1/150	Out > In p < 0.0001 BF < 1/150	p = 0.92 BF = 7.47	Out > In p < 0.0001 BF < 1/150
	Overall recall (In vs. Out)	p = 0.21 BF = 3.46	p = 0.56 BF = 6.35	p = 0.14 BF = 2.56	In > Out p = 0.03 BF = 0.96
lag-CRP	Lag (1 to 29)	Nearby > far away p < 0.0001 BF < 1/150	Nearby > far away p < 0.0001 BF < 1/150	Nearby > far away p < 0.0001 BF < 1/150	Nearby > far away p < 0.0001 BF < 1/150
	Direction (Forward vs. Backward)	p = 0.79 BF = 37.41	Forward > backward p = 0.02 BF = 1.96	p = 0.80 BF = 38.88	p = 0.20 BF = 24.95
	Attentional State (In vs. Out)	p = 0.74 BF = 37.92	p = 0.57 BF = 31.22	p = 0.20 BF = 12.39	In > Out p = 0.0485 BF = 2.08
	Lag x Direction	p = 0.33 BF > 150	Forward > backward for nearby > far away p = 0.003 BF = 0.56	Forward > backward for nearby > far away p = 0.007 BF = 18.22	Forward > backward for nearby > far away p = 0.02 BF > 150
	Lag x Attentional State	p = 0.64 BF > 150	p = 0.39 BF > 150	p = 0.66 BF > 150	p = 0.84 BF > 150
	Direction x Attentional State	p = 0.33 BF = 20.99	p = 0.79 BF = 26.88	p = 0.73 BF = 26.99	p = 0.47 BF = 29.46
	Lag x Direction x Attentional State	p = 0.29 BF > 150	p = 0.19 BF > 150	p = 0.51 BF > 150	p = 0.90 BF > 150
lag-CRP +/- 1	Lag (+1 vs. -1)	+1 > -1 p = 0.043 BF = 0.52	+1 > -1 p = 0.0009 BF = 0.024	+1 > -1 p = 0.0007 BF = 0.028	+1 > -1 p = 0.002 BF = 0.043
	Attentional State (In vs. Out)	p = 0.22 BF = 4.25	p = 0.66 BF = 6.89	p = 0.46 BF = 6.36	p = 0.59 BF = 8.96
	Lag x Attentional State	p = 0.76 BF = 5.09	p = 0.26 BF = 2.86	p = 0.26 BF = 3.08	p = 0.92 BF = 7.20
Event Analysis	Transition Type	p < 0.0001 BF < 1/150	p < 0.0001 BF < 1/150	p < 0.0001 BF < 1/150	p < 0.0001 BF < 1/150
	Attentional State (In vs. Out)	p = 0.85 BF = 8.87	p = 0.54 BF = 7.77	p = 0.13 BF = 4.41	p = 0.27 BF = 6.93
	Transition Type x Attentional State	p = 0.83 BF = 19.13	p = 0.53 BF = 15.08	p = 0.26 BF = 10.66	p = 0.58 BF = 25.54

(Rouder et al., 2012).

In addition to Bayes Factors for the effects of interest, we also included Bayesian model comparisons for the lag-CRP and event segment analyses. Here, we compared a null model that did not include attentional state as a factor (whether as a main effect or interaction term) to a full model that included attentional state and its interactions (Kruschke, 2011; Liu & Aitkin, 2008). The Bayes Factor for the null model was then divided by the Bayes Factor for the full model (both computed using the function lmBF) to obtain evidence for the null (i.e., values larger than 3 indicate evidence for the null model).

We also provide traditional null-hypothesis significance testing throughout the paper (i.e., determining whether results are statistically significant or not based on whether *p*-values are <0.05). In almost all cases, p-values and Bayes Factors are in agreement. There are, however, some situations in which p-values are <0.05 but evidence from Bayes Factors does not strongly support the alternative hypothesis; this only happens in rare cases and does not affect our main conclusions. In all analyses testing our main hypotheses — with respect to the temporal organization of recall based on attentional state at encoding — interpretation of p-values and Bayes Factors leads to the same conclusion.

2.2. Results

2.2.1. Defining attentional states at encoding

In the encoding task, participants viewed images of objects and judged each as being larger or smaller than a shoebox. Overall, mean response time (RT; defined from the onset of an image fading in; see Section 2.1: **Methods** and Fig. 1) was 3.94 (SD = 0.58 s). Median RT was 4.00s.

We defined attentional states by performing a variance time course (VTC) analysis on the encoding phase RTs (Esterman et al., 2013). This procedure identifies two attentional states: the good "in the zone" attentional state (trials with lower RT variability, i.e., RTs closer to the mean) and the worse "out of the zone" attentional state (trials with higher RT variability, i.e., RTs farther away from the mean). Fig. 3A shows the VTC analysis for one sample participant.

The mean length of a continuous "in the zone" segment was 3.59 trials (SD = 0.46) and the mean length of a continuous "out of the zone" segment was 3.59 trials (SD = 0.53; Note that each trial was 6 s long). The mean number of fluctuations within a block (i.e., the number of times participants transitioned from one state to another) was 7.38 (SD = 1.04). The number of trials within a continuous segment ranged from 1 to 14, for both "in the zone" and "out of the zone" states, across all blocks and participants.

As an initial validation check of the VTC approach, we examined

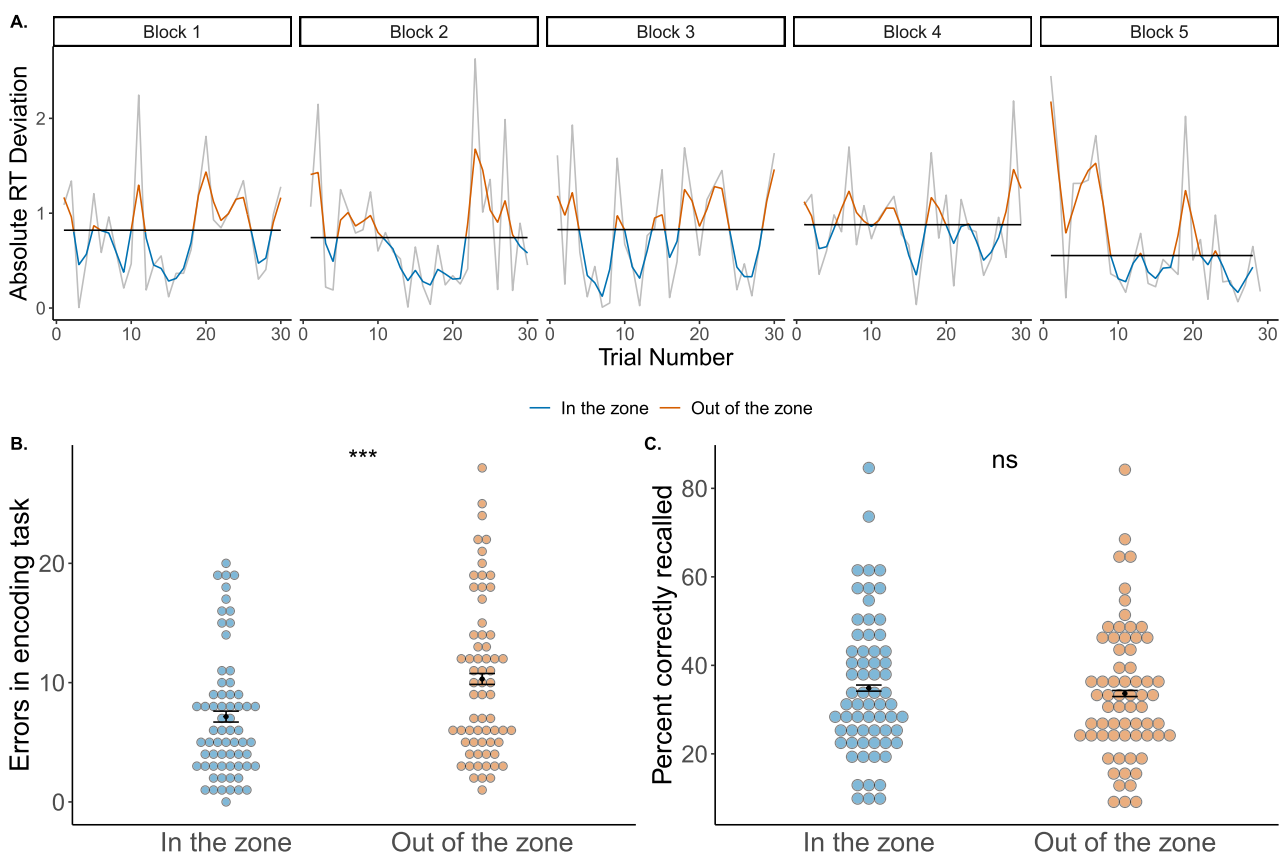

Fig. 3. Encoding task performance and recall performance in Study 1. Encoding errors differ between attentional states but recall performance does not. **A.** Variance Time Course (VTC) analysis for a sample participant. Two attentional states were identified by a median split of smoothed, absolute RT deviations from the mean: 1) an "in the zone" attentional state (blue) with lower RT variability, i.e., RTs closer to the mean and, 2) an "out of the zone" attentional state (orange) with higher RT variability, i.e., RTs farther away from the mean. Horizontal black lines indicate the median absolute RT deviation per block. Gray curves indicate raw (unsmoothed) RT deviation per block. **B.** Individual points show the number of encoding judgement errors made by each participant during "in the zone" and "out of the zone" attentional states. Participants made significantly more encoding errors during the "out of the zone" state. **C.** Individual points show the percentage of items correctly recalled by each participant as a function of whether items were encoded "in the zone" or "out of the zone". There was no difference in recall performance between the two states. Black points in panels B & C indicate the mean of the measure; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". *** $p < 0.0001$. ns = not statistically significant. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

whether RT variability changed over the course of the experiment, as would be predicted if participants gradually lost focus or became fatigued over time. Indeed, RTs became progressively more variable across blocks ($F_{3.49,223.21} = 4.91$, $p = 0.0015$, $\eta_p^2 = 0.07$), consistent with the idea that more variable RTs may index worsening attention and differentiate "in the zone" from "out of the zone" states.

These attentional states were used to examine accuracy on the encoding task and subsequent recall performance, described below.

2.2.2. More encoding errors during "out of the zone" attentional states

Participants generally performed very well on the encoding task ("Is this object smaller or larger than a shoebox?"). The mean percentage of trials participants responded to was 97.82% (SD = 3.11%, Median = 99.33%). Mean accuracy (as defined in Section 2.1: **Methods**) was 88.66% (SD = 7.12%, Median = 91.33%).

We next examined errors in the encoding task as a function of attentional state. Prior studies have shown that "out of the zone", vs. "in the zone", attentional states are associated with more errors (Esterman et al., 2013; Rosenberg et al., 2013). We replicated these findings in the current study. A paired samples *t*-test revealed that participants made significantly more errors in the encoding task during an "out of the zone" attentional state (mean ± SD: 10.31 ± 6.67) compared to an "in the zone" attentional state (7.15 ± 5.32; $t_{64} = 4.87$, $p < 0.0001$, Cohen's dz. = 0.60, 95% CI [1.86, 4.45], BF < 1/150, Fig. 3B). Thus, the VTC analysis is successful in identifying fluctuations between better and worse attentional states, even in our modified procedure.

As an additional validation check of whether RTs are sensitive to attentional fluctuations in our task, we compared RTs before vs. after an error, as done in prior work (e.g., Dutilh et al., 2012; Fortenbaugh et al., 2015). To that end, we used a linear mixed-effects model to predict RT from trial position (pre vs. post error), with participant-level random effects for the intercept (following the procedure used in Decker et al., 2020). We found reliable post-error speeding: RTs after an error were significantly faster than those before an error ($\beta = -230.82$, S.E. = 74.15, $t_{1955} = -3.11$, $p = 0.0019$, 95% CI [−376.15, −85.49]). Although counterintuitive, several other studies have found post-error speeding (Damaso, Williams, & Heathcote, 2020; Notebaert et al., 2009; Purcell & Kiani, 2016; Williams, Heathcote, Nesbitt, & Eidels, 2016); this may occur if participants occasionally "zone out" or mind wander, leading to inefficient processing of an item on a given trial and subsequent speeding once participants catch themselves being off task. Thus, this analysis provides additional evidence that RTs can index attention in our task.

2.2.3. Overall recall does not differ between the two attentional states

We next turned to examining memory for the objects viewed during the encoding task. Mean recall accuracy, calculated as the percentage of items correctly recalled across all blocks, was 34.28% (SD = 14.67).

We then separately examined recall performance based on whether items were encoded "in the zone" or "out of the zone". We hypothesized that recall performance would be superior for "in the zone" attentional states. However, we did not find a significant difference in recall performance between "in the zone" (Mean ± SD: 34.86% ± 15.70) and "out of the zone" (33.64% ± 15.20) attentional states ($t_{64} = 1.26$, $p = 0.21$, Cohen's dz. = 0.16, 95% CI [−0.72, 3.17], BF = 3.46, Fig. 3C). Thus, even though these attentional states differed in *online* task performance, subsequent recall was surprisingly not different.

2.2.4. No differences in temporal contiguity or forward asymmetry between the two attentional states

Although overall recall was not different between "in the zone" and "out of the zone" attentional states, it is possible that there may be subtle differences in *how* information is recalled. We therefore turned to our main analyses of interest, which explore the temporal organization of recall. We hypothesized that "in the zone" attentional states are more conducive to maintaining temporal context representations that facilitate temporally organized recall. We therefore used lag-CRP curves to test whether temporal contiguity and/or forward asymmetry were stronger for "in the zone" vs. "out of the zone" states.

Fig. 4A shows the overall lag-CRP curve, across participants and blocks, regardless of attentional state at encoding (see Table 2 for analyses of the overall lag-CRP curve); Fig. 4B shows the lag-CRP curves, across participants and blocks, separately for items that were encoded "in the zone" and "out of the zone". These curves depict the probability of recalling two items successively based on their relative position, or lag, at encoding. Note that, across all of our Studies, the CRP values at nearby lags are lower than what is typically observed in many recall studies (e.g., Healey & Kahana, 2014; Kahana, 1996; Sederberg et al., 2010). This is likely because the lists used in the current set of Studies are longer than those of most list-learning recall studies and, concomitantly, more recall transitions are available for each item. Thus, recall transitions could be spread across more lags. This is consistent with the finding that lag-CRP values at nearby lags tend to be lower for longer lists (Healey et al., 2019).

We conducted a three-way repeated-measures ANOVA on the lag-CRP measures (i.e., lag-conditional recall shown in Fig. 4B) with attentional state ("in the zone" vs. "out of the zone"), absolute lag (1 to 29), and direction (forward vs. backward) as factors (results summarized in Table 3). We expected that we might find 1) an interaction between attentional state and absolute lag, indicating that nearby recall transitions would be more likely for items encoded "in the zone" vs. "out of the zone" and 2) an interaction between attentional state and direction, indicating a stronger forward asymmetry bias for items encoded "in the zone". We found a significant main effect of absolute lag ($F_{9.07,\ 580.52} = 9.60$, $p < 0.0001$, $\eta_p^2 = 0.13$, BF < 1/150): during recall, individuals were more likely to transition to items that were encoded nearby vs. farther away. We did not find a significant main effect of direction ($F_{1,64} = 0.074$, $p = 0.79$, $\eta_p^2 = 0.001$, BF = 37.41) nor an interaction between direction and absolute lag ($F_{8.44,\ 540.05} = 1.15$, $p = 0.33$, $\eta_p^2 = 0.02$, BF > 150).

We next examined main effects and interactions involving attentional state. There was no main effect of attentional state ($F_{1,64} = 0.11$, $p = 0.74$, $\eta_p^2 = 0.002$, BF = 37.92). Additionally, there was no interaction between attentional state and direction ($F_{1,64} = 0.96$, $p = 0.33$, $\eta_p^2 = 0.015$, BF = 20.99), no interaction between attentional state and absolute lag ($F_{7.66,490.37} = 0.75$, $p = 0.64$, $\eta_p^2 = 0.012$, BF > 150), nor a three-way interaction between absolute lag, direction, and attentional state ($F_{8.27,\ 529.45} = 1.21$, $p = 0.29$, $\eta_p^2 = 0.02$, BF > 150). Hence, there was no significant difference between the attentional states in temporal contiguity or forward asymmetry.

We next examined forward asymmetry differences between the two attentional states at the nearby lags of ±1. This was done to determine whether the lack of a forward asymmetry difference between the attentional states was due to inclusion of all lags in the CRP curves: forward symmetry is sometimes most pronounced for nearby lags. From a two-way repeated-measures ANOVA with lag (+1 vs. -1) and attentional state ("in the zone" vs. "out of the zone") as factors, we did not find an effect of attentional state ($F_{1,64} = 1.51$, $p = 0.22$, $\eta_p^2 = 0.023$, BF = 4.25) nor an interaction between attentional state and lag ($F_{1,64} = 0.091$, $p = 0.76$, $\eta_p^2 = 0.001$, BF = 5.09).

Finally, we conducted a Bayesian model comparison in which we compared the three-way ANOVA reported above (lag x direction x attentional state) to a null model without an attentional state variable (lag x direction). This analysis revealed very strong evidence in favor of the null model (BF > 150).

All together, across these lag-CRP analyses, we did not find any differences in the temporal organization of recall for items encoded "in the zone" vs. "out of the zone".

2.2.5. No differences in event transition types between the two attentional states

The lag-CRP analyses above examined temporal organization

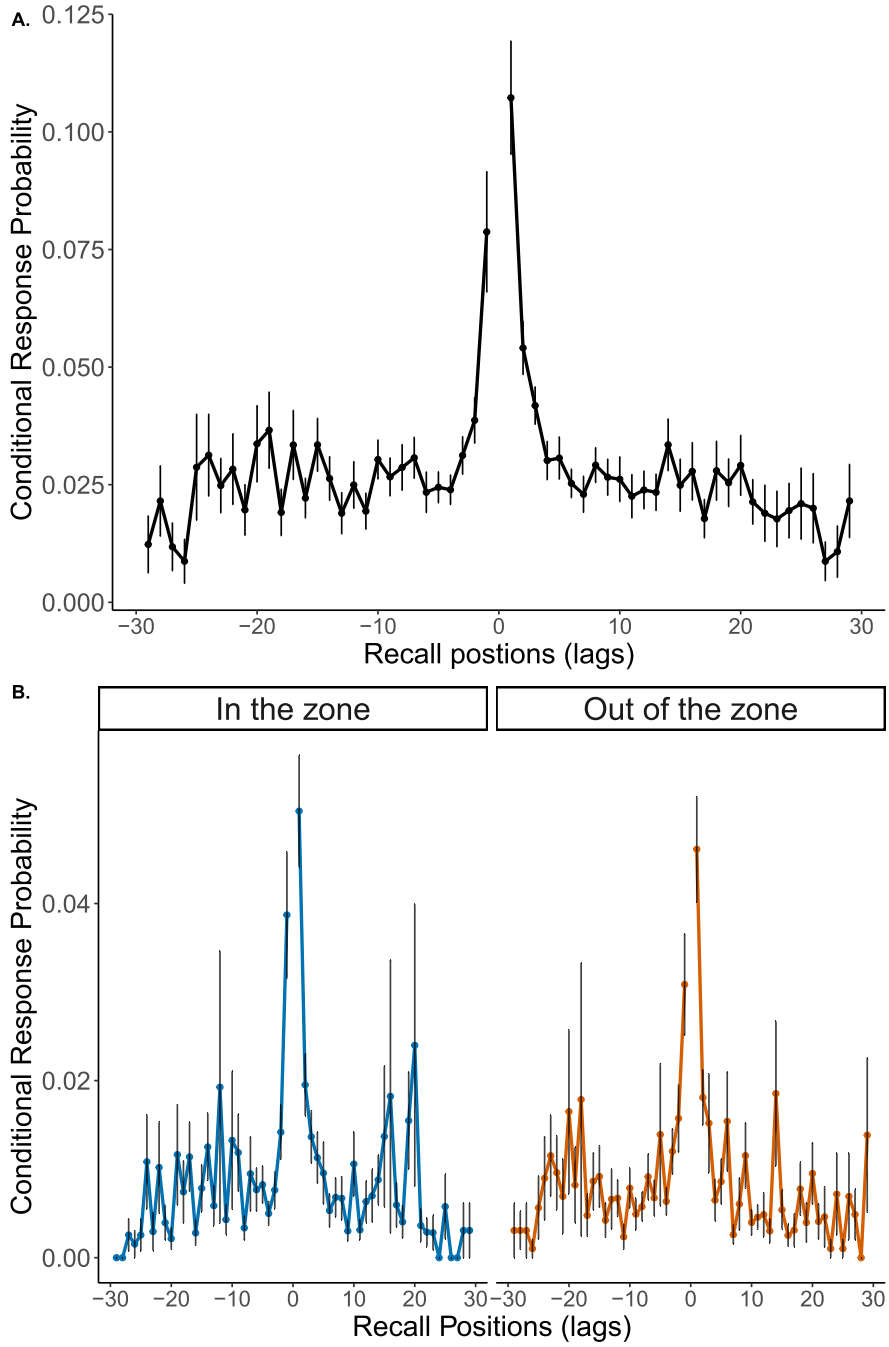

Fig. 4. Lag-CRP curves overall and by attentional state for Study 1. **A.** Overall lag-CRP curve across participants and blocks. **B.** Lag-CRP curves plotted separately for items encoded "in the zone" (left) and "out of the zone" (right). There was no difference between the two attentional states in the temporal contiguity or forward asymmetry of recall. Error bars represent the standard error.

differences between the two attentional states, but they do not take into account qualitatively different types of transitions that could occur within an attentional state (such as transitions between different event segments) or transitions from one state to another (see Fig. 2). However, it is possible that attentional fluctuations act in a similar way to event boundaries (DuBrow & Davachi, 2013, 2016; Ezzyat & Davachi, 2010; Heusser et al., 2018), such that recall is more temporally clustered within segments than across segments, and may occasionally "leap" between segments of a similar cognitive state (Chan et al., 2017).

We therefore examined recall as a function of the type of transition (Fig. 2). If "in the zone" (vs. "out of the zone") attentional states are more conducive to maintaining a temporal context representation, and this temporal context representation is reinstated every time an individual is "in the zone", two predictions could be made. First, that recall transitions within an "event segment" may be more likely for items encoded "in the zone" vs. "out of the zone" (i.e., same state, same event transitions; Fig. 2), and second, that recall "leaps" to a different event segment in the same attentional state may be more likely for items encoded "in the zone" vs. "out of the zone" (i.e., same state, different event transitions).

To test this, for each type of transition ("same state, same event"; "same state, different event"; "different state, different event"), we calculated the number of transitions made during recall divided by the number of opportunities to make such transitions (see Section 2.1.2.5:

Recall Transitions by Event Segment). We performed a two-way repeated-measures ANOVA with transition type (3 levels) and attentional state ("in the zone" vs. "out of the zone") as factors. We expected to find an interaction between attentional state and transition type. Specifically, we expected to find more "same state, same event" and "same state, different event" transitions for "in the zone" vs. "out of the zone" attentional states.

We found a main effect of transition type ($F_{1.07,\ 68.51} = 32.24$, $p < 0.0001$, $\eta_p^2 = 0.33$, BF < 1/150). There was no main effect of attentional state ($F_{1,64} = 0.037$, $p = 0.85$, $\eta_p^2 = 0.0006$, BF = 8.87) nor a significant interaction between attentional state and transition type ($F_{1.28,\ 81.66} = 0.086$, $p = 0.83$, $\eta_p^2 = 0.001$, BF = 19.13).

Given the main effect of transition type, we conducted follow-up *t*-tests to compare them, collapsing across the two attentional states. We found that "same state, same event" transitions were significantly more likely than "same state, different event" transitions ($t_{64} = 5.77$, $p < 0.0001$, Cohen's dz. = 0.72, 95% CI [0.013, 0.027]) and significantly more likely than "different state, different event" transitions ($t_{64} = 5.71$, $p < 0.0001$, Cohen's dz. = 0.71, 95% CI [0.012, 0.025]). "Different state, different event" transitions were also more likely compared to "same state, different event" transitions ($t_{64} = 2.13$, $p = 0.04$, Cohen's dz. = 0.26, 95% CI [0.00010, 0.003]).

This pattern of results (Fig. 5) is consistent with the temporal contiguity effect: participants are more likely to make recall transitions to nearby items. In particular, items encoded in the same state and same event are the closest to one another ("same state, same event" transitions; e.g., recall transitions within Event 1 in Fig. 2), and are more likely to be recalled together than items that span events ("same state, different event" and "different state, different event" transitions in Fig. 2).

As for the lag-CRP analysis, we conducted a Bayesian model comparison in which we compared the two-way ANOVA reported above (transition type x attentional state) to a null model without an attentional state variable (one-way ANOVA with transition type). This analysis revealed very strong evidence in favor of the null model (BF > 150). Therefore, across these analyses, we found no evidence that recall transitions across different event types varied based on attentional state at encoding.

2.2.6. Secondary analysis: faster vs. slower RTs

In addition to defining attentional states based on the VTC approach (in which RTs that are either too fast or too slow are considered to reflect a poor attentional state), we conducted a secondary analysis in which we divided attentional states into those characterized by faster vs. slower RTs, in line with other research that makes this distinction (e.g., deBettencourt et al., 2018; Decker et al., 2020; Wakeland-Hart et al., 2022). We therefore analyzed the data the same way as we did for the VTC analysis (including, for example, interpolating RTs on incorrect trials), except that trials were divided by a median split into those associated with faster vs. slower RTs (rather than RTs closer vs. further from the mean).

There were no differences in online errors during attentional states characterized by faster vs. slower RTs ($t_{64} = 1.90$, $p = 0.06$, Cohen's dz. = 0.24, 95% CI [−0.05, 2.05], BF = 1.36). Note that the numerical direction of this effect (more errors for attentional states associated with slower vs. faster RTs) is the opposite from that expected based on some similar prior work, in which more errors were made following faster vs. slower RTs (deBettencourt et al., 2018; Wakeland-Hart et al., 2022). It is, however, consistent with findings that longer eye fixations and slower response times are associated with mind wandering and attentional lapses (e.g., Smallwood et al., 2003; Weissman et al., 2006; Feng et al., 2013; Yanko & Spalek, 2013; Kam & Handy, 2014; Henríquez et al., 2016; Krasich et al., 2018; Zhang et al., 2021).

Subsequent recall was not different between attentional states associated with faster vs. slower RTs at encoding ($t_{64} = 1.21$, $p = 0.23$, Cohen's dz. = 0.16, 95% CI [−0.77, 3.15], BF = 3.66). The lag-CRP analysis revealed only a main effect of absolute lag (nearby > far away; $F_{12.68,\ 811.72} = 7.96$, $p < 0.0001$, $\eta_p^2 = 0.11$, BF < 1/150); all other main effects and interactions were not statistically significant (all *ps* > 0.26; all other BFs > 20.40).

The event segment analysis revealed a main effect of transition type ($F_{1.04,\ 66.24} = 35.41$, $p < 0.0001$, $\eta_p^2 = 0.36$, BF < 1/150). Follow-up *t*-tests showed significantly more "same state, same event" transitions compared to "same state, different event" transitions ($t_{64} = 5.90$, $p < 0.0001$, Cohen's dz. = 0.73, 95% CI [0.01, 0.03]), and "different state, different event" transitions ($t_{64} = 6.10$, $p < 0.0001$, Cohen's dz. = 0.76, 95% CI [0.015, 0.029]). "Different state, different event" transitions were also more likely than "same state, different event" transitions (t_{64}

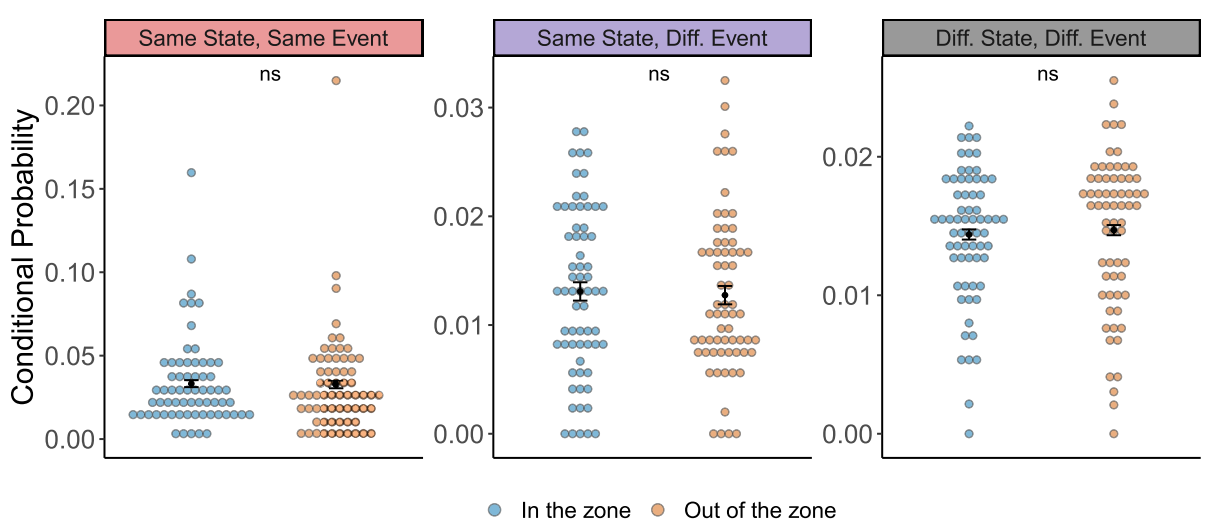

Fig. 5. Recall transitions as a function of event type for Study 1. Recall transitions are shown based on whether they occurred within an "event segment" of a particular attentional state (same state, same event), across event segments of a given attentional state (same state, different event), or between attentional states (different state, different event; see Fig. 2). There were no significant differences between the two attentional states in any transition type. Individual points indicate the conditional probability of each transition type (i.e., the number of times each transition type occurred divided by the number of opportunities to make a transition of that type) for each individual, separately for items encoded "in the zone" and "out of the zone". Black dots indicate the mean; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". ns = not statistically significant.

$= 2.03$, $p = 0.047$, Cohen's dz. $= 0.25$, 95% CI [0.00002, 0.003]). There was no main effect of attentional state ($F_{1,64} = 0.25$, $p = 0.62$, $\eta_p^2 = 0.004$, BF $= 8.50$) nor an attentional state by transition type interaction ($F_{1.12,\ 71.44} = 1.24$, $p = 0.28$, $\eta_p^2 = 0.02$, BF $= 10.35$).

A Bayesian model comparison for the lag-CRP analysis, which compared the above model to a null model without an attentional state variable, revealed very strong evidence in favor of the null model (BF $>$ 150). The same approach for the event segment analysis revealed strong evidence in favor of the null model (BF $= 86.32$).

Thus, this alternative characterization of attentional states also failed to reveal differences in the temporal organization of recall. This alternative approach is, however, more difficult to interpret than the VTC analysis: unlike the VTC analysis, which was successful in detecting differences in online performance for "in the zone" vs. "out of the zone" attentional states, this alternative approach did not reliably predict online errors. Thus, this alternative approach may not be as powerful in identifying fluctuating attentional states.

2.3. Discussion

We hypothesized that "in the zone" attentional states, vs. "out of the zone" states, are more conducive to maintaining temporal context representations that can facilitate temporally organized recall. We found no support for this hypothesis. There was no difference between attentional states in overall recall; there was also no difference in forward asymmetry or temporal contiguity as assessed by the lag-CRP curves; and finally, there was no difference in the types of event transitions made in recall.

We replicated prior results in showing more online errors (in the encoding task) during "out of the zone" vs. "in the zone" attentional states. Yet, subsequent recall was not different overall between the two attentional states, and no measures of recall organization showed a difference. One possibility is that our task only yielded moderate attentional fluctuations, which were not strong enough to produce effects on recall. This may be because our task deviated from the original gradCPT in several ways (Esterman et al., 2013; Rosenberg et al., 2013). The original gradCPT requires a habitual response to frequent trials and a withheld response to infrequent trials. Here, we had participants provide a binary judgement (using one of two keys) on each trial, for which the responses were similar in frequency. Thus, the traditional manipulation might be more effective in inducing "out of the zone" states because 1) the same response is made most of the time and 2) the judgement used in the current task might be more difficult and subjective, and thus may have required more focused attention.

Furthermore, the subjectivity of the judgements used in this Study may have made it difficult to separate "in the zone" and "out of the zone" states. This may have occurred if RTs to ambiguous images (e.g., cowboy hat, soda bottle) were slow, as participants considered how either response (bigger or smaller than a shoebox) could be justified. If that was the case, responses to ambiguous items may be incorrectly considered "out of the zone", obscuring true differences between attentional states. We conducted two additional analyses to test this possibility. First, we examined if RTs to ambiguous items were slower than those to unambiguous items; that was indeed the case ($t_{64} = 3.88$, $p = 0.0003$, Cohen's dz. $= 0.48$, 95% CI [-0.34, -0.11]). We next repeated all of our analyses after interpolating RTs for ambiguous items, using the same procedure used for error trials. The pattern of results was unchanged: the statistically significant results were in online errors ("out of the zone" $>$ "in the zone"; $t_{64} = 3.85$, p $= 0.0003$, Cohen's dz. $= 0.48$, 95% CI [1.20, 3.79]); the main effect of absolute lag in the lag-CRP analysis (nearby transitions $>$ far away transitions; $F_{9.87,\ 631.61} = 8.49$, $p < 0.0001$, $\eta_p^2 = 0.12$; all other $ps > 0.22$); and the main effect of transition type in the event segment analysis (same pattern as reported in the main analysis; $F_{1.08,68.92} = 29.59$, p < 0.0001, $\eta_p^2 = 0.32$; all other $ps > 0.45$). Thus, inclusion of ambiguous items had no detectable effect on our analyses of interest.

A final difference between our Study and prior work is that our blocks were relatively short: 3 min relative to 8 min in a traditional gradCPT. These short blocks may not have induced strong enough attentional fluctuations to see large effects on recall organization. In Study 2, we modified our task to address these limitations.

3. Study 2

3.1. Overview

In Study 2, we sought to address limitations of Study 1 that may have made attentional fluctuations relatively weak and thus, limited our chance of seeing strong effects on subsequent recall. First, we changed our task to be aligned with the traditional gradCPT by using a go/no-go approach. Most trials were "go" trials in which a participant made a response to non-food items. On a minority of trials ("no-go"), which occurred 10% of the time, a food item was presented and participants had to withhold their response. This approach should make the "go" response habitual, making it more likely that individuals will "zone out" due to the repetitive nature of the task. We did not include any stimuli for which a "food" vs. "not food" judgement would be ambiguous (e.g., animals); this addresses a limitation of Study 1, in which some ambiguous images were used. Second, we made our encoding blocks longer (3 blocks of 8 min, rather than 5 blocks of 3 min in Study 1). This was done with the hope of encouraging stronger periods of "zoning out". To that end, we increased our block length by adding more items per block: 80 items per block rather than the 30 in Study 1. A small pilot study indicated that participants recalled fewer words with the longer blocks; we therefore added a block beyond what was needed to (roughly) balance the total number of items in Study 1 and Study 2 (i.e., 3 blocks, rather than 2, of 80 items each). This allowed us to match the average number of words recalled across Study 1 and Study 2.

3.2. Methods

3.2.1. Design

3.2.1.1. Participants. We report data from 68 participants ($M_{age} = 22.62 \pm 5.15$, $M_{education} = 13.99 \pm 1.65$; see Table 1 for demographics). We do not report data from an additional 15 participants, who were excluded due to image loading errors ($N = 3$), low response rate during the encoding task ($<$80%, $N = 7$), outlier response accuracy during the encoding task ($>$3 SD from the group mean; $N = 1$), and recall recording issues ($N = 4$). Of the final sample, 48 participants were recruited from the Columbia University participant pool and the rest (20 participants) were recruited through Prolific (www.prolific.co). All participants completed an online version of the task hosted on the Gorilla platform (www.gorilla.sc; Anwyl-Irvine et al., 2020). Informed consent was obtained in accordance with the Columbia University Institutional Review Board.

3.2.1.2. Stimuli. Stimuli were identical to Study 1 with the following exceptions. We chose 240 images instead of 191 from the pre-curated databases. Color images were converted to grayscale. 90% of the images (216 images) were non-food (i.e., inedible) items and 10% (24 images) were food items. No ambiguous stimuli were included (e.g., animals). The 240 images were divided into 3 lists of 80 images each (8 food, 72 non-food images). For this and subsequent Studies, the stimulus lists were created by manually ensuring there were an equal number of items from a category (e.g., tools, furniture) in each list. This manual sorting was used instead of the OptSeg algorithm from Study 1 because there were fewer lists, and it was thus more tractable to do the balancing manually (OptSeg was relatively slow to run and required manual checking of the generated lists; thus, it was more efficient to make the lists manually).

3.2.1.3. Procedure. The procedure was identical to Study 1 with the following exceptions. The experiment consisted of 3 blocks, each of which included a study phase, a distractor phase, and a recall phase (Fig. 1). In each study phase, participants viewed 80 trial-unique items, which transitioned slowly from one into another as in Study 1. For each presented image, participants were asked to judge if the depicted object was "a food or non-food item". Importantly, they were asked to press a button when it was a non-food item (the dominant category), but withhold their response when it was a food item. This change aligned our task with the traditional gradCPT, such that participants habitually pressed one response 90% of the time, which may make it more likely for them to "zone out".

The distractor phase was identical to that in Study 1. The recall phase was similar to Study 1. Participants were initially given 2 min to verbally free recall items from the study phase. Unlike Study 1, after the initial 2 min of recording (the maximum allowed on Gorilla), participants were given the option to begin recording for another 2 min if they wanted to recall more objects. This was done because the blocks in Study 2 were longer than those in Study 1; thus, we wanted to give participants more time for free recall if they needed it.

Participants did not perform a practice block before beginning the task blocks, but were given video instructions on how to perform the task.

3.2.2. Analyses

3.2.2.1. Defining attentional states at encoding. Attentional states were defined using RTs for the judgements made by participants during the study phase, in a procedure similar to Study 1. Unlike Study 1, however, participants were supposed to withhold responses on some trials; thus, some correct responses did not have an associated RT. Therefore, we first calculated — for correct trials with a response — the absolute deviation of the trial RT from the within-block mean, as in Study 1. Next, RT deviations for trials without a response (whether correctly withheld on "no-go" food trials or incorrectly withheld on "go" non-food trials) and trials with an incorrect response were interpolated from the two surrounding trials, as done in other studies employing this method (Esterman et al., 2013; Rosenberg et al., 2013). All other steps were performed in an identical manner to Study 1. This resulted in trials being divided into two attentional states: "in the zone" states with lower RT variability (i.e., RTs closer to the mean) and "out of the zone" states with higher RT variability (i.e., RTs farther away from the mean; Fig. 6A).

3.2.2.2. Errors at encoding. Similar to Study 1, we first sought to replicate the finding that "out of the zone" (vs. "in the zone") attentional states are associated with more errors (Esterman et al., 2013; Rosenberg et al., 2013). Errors in the encoding task were calculated as the sum of the number of incorrect button presses to a "no-go" food trial (commission errors) and the number of failures to respond to a "go" non-food

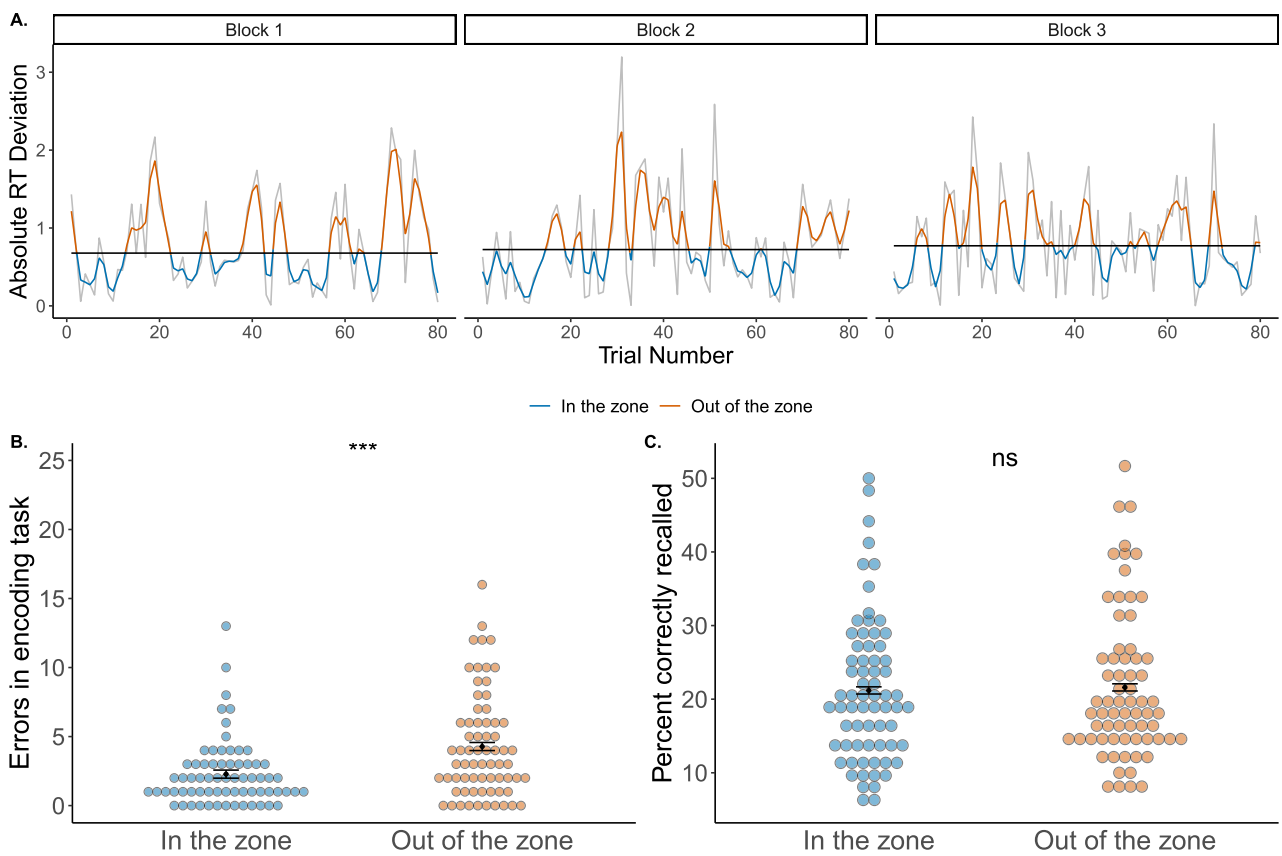

Fig. 6. Encoding task performance and recall performance in Study 2. Encoding errors differ between attentional states but recall performance does not. **A.** Variance Time Course (VTC) analysis for a sample participant, depicting "in the zone" (blue) and "out of the zone" (orange) attentional states. Horizontal black lines indicate the median absolute RT deviation per block. Gray curves indicate raw (unsmoothed) RT deviation per block. **B.** Individual points show the number of encoding judgement errors made by each participant during "in the zone" and "out of the zone" attentional states. Participants made significantly more encoding errors during the "out of the zone" state. **C.** Individual points show the percentage of items correctly recalled by each participant as a function of whether items were encoded "in the zone" or "out of the zone". There was no difference in recall performance between the two states. Black points in panels B & C indicate the mean of the measure; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". *** $p < 0.0001$. ns = not statistically significant. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

trial (omission errors). We examined whether the number of errors made during "out of the zone" attentional states was higher than the number of errors made during "in the zone" states. Group-level analyses were conducted with a paired-samples *t*-test.

3.2.2.3. Recall performance. Recall performance and related analyses were identical to those in Study 1.

3.2.2.4. Temporal organization of recall. Analyses of temporal organization of recall were identical to those in Study 1. Note that in Study 1, actual and possible transitions ranged from −29 to +29 and the entire range was used in lag-CRP analyses. In Study 2, the range of actual and possible transitions is −79 to +79, because the length of the encoding list is 80 items. However, for the lag-CRP analyses of interest, we only used actual and possible transitions between −29 to +29. There were two reasons for this: First, transitions at the farther lags were rare (26 trials or fewer, across all 3 blocks for all 68 participants combined, for a given lag further away than ±29), and hence the lag-CRP estimates were particularly noisy at those lags. Second, to facilitate comparison across studies, we opted to keep our analyses consistent by using the range from Study 1.

3.2.2.5. Recall transitions by event segment. Analyses of recall transitions by event segment were identical to those in Study 1.

3.3. Results

3.3.1. Defining attentional states at encoding

In the encoding task, participants viewed images of objects and judged each as being a non-food item (with a button press) or a food item (by withholding a response). Overall, mean RT (defined from the onset of an image fading in; see **Section 2.1**: **Methods**) was 3.74 s (SD = 0.72). Median RT was 3.77 s.

As before, we defined "in the zone" and "out of the zone" attentional states by performing a variance time course (VTC) analysis on the encoding phase RTs (see **Section 3.1**: **Methods**). Fig. 6A shows the VTC analysis for one sample participant in Study 2.

The mean length of an "in the zone" segment was 3.95 trials (SD = 0.42) and the mean length of an "out of the zone" segment was 3.90 trials (SD = 0.43; Note that each trial was 6 s long). The mean number of fluctuations within a block (i.e., the number of times participants transitioned from one state to another) was 19.47 (SD = 2.25). The number of trials within a segment ranged from 1 to 27 for "in the zone" states and from 1 to 23 for "out of the zone" states, across all blocks and participants.

As for Study 1, we sought an initial validation of the VTC approach by testing whether RTs were increasingly variable over the experimental session. Once again, RTs became progressively more variable over blocks ($F_{1.77,118.34} = 5.66$, $p = 0.006$, $\eta_p^2 = 0.08$), as individuals presumably became more fatigued and less focused. This lends support to the use of RT variability to differentiate "in the zone" and "out of the zone" states.

These attentional states were used to examine accuracy on the encoding task and subsequent recall performance, described below.

3.3.2. More encoding errors during "out of the zone" attentional states

Participants performed very well on the encoding task ("Is this object food or a non-food item?"). They responded to 99.41% (SD = 2.57%; Median = 100%) of the "go" non-food trials, which required a response. Mean accuracy (including correct responses on "go" trials and withheld responses on "no-go" trials) was 96.87% (SD = 2.74%; Median = 97.92%).

We next examined errors in the encoding task as a function of attentional state. We replicated prior studies (Esterman et al., 2013; Rosenberg et al., 2013) and Study 1: A paired samples *t*-test revealed that participants made significantly more encoding errors during an "out of the zone" attentional state (mean ± SD: 4.28 ± 3.78) compared to an "in the zone" attentional state (2.28 ± 2.43; $t_{67} = 4.83$, $p < 0.0001$, Cohen's dz. = 0.59, 95% CI [1.17, 2.83], BF < 1/150, Fig. 6B). Thus, the VTC analysis remains successful in identifying fluctuations between better and worse attentional states.

As in Study 1, we performed an additional validation check of whether RTs are sensitive to attentional fluctuations in our task, by comparing RTs before vs. after an error. Unlike Study 1 (for which participants should respond on every trial), there were two types of potential errors in Study 2: omission errors, for which participants fail to make a button response when they should, and commission errors, for which participants make a button response when they should have withheld it. We therefore examined RTs before vs. after these error types.

Participants were faster to respond after vs. before an omission error ($\beta = -1148.6$, S.E. = 152.00, $t_{310.40} = -7.56$, $p < 0.0001$, 95% CI [−1446.53, −850.60]). This is consistent with participants "zoning out" and failing to respond to an image, and subsequently speeding up once they realize that a trial was missed. Conversely, participants responded slower after vs. before a commission error ($\beta = 392.81$, S.E. = 122.39, $t_{330.94} = 3.21$, $p = 0.002$, 95% CI [152.94, 632.70]). This is consistent with fast, habitual responses leading to errors, with post-error slowing once an erroneous response has been detected. Together, post-error slowing after commission errors and post-error speeding after omission errors show that RTs can index attentional fluctuations in our task.

3.3.3. Overall recall does not differ between the two attentional states

We next turned to examining memory for the objects encoded during the study phase. Mean recall was 21.38% (SD = 9.71).

We then separately examined recall based on whether items were encoded "in the zone" or "out of the zone". As in Study 1, we did not find a significant difference in recall performance between "in the zone" (Mean ± SD: 21.19% ± 9.91) and "out of the zone" (21.60% ± 10.18) attentional states ($t_{67} = 0.59$, $p = 0.56$, Cohen's dz. = 0.072, 95% CI [−0.97, 1.79], BF = 6.35, Fig. 6C). Thus, as in Study 1, these states differed in performance during the encoding task but showed no differences in subsequent recall.

3.3.4. No differences in temporal contiguity or forward asymmetry between the two attentional states

As in Study 1, we constructed separate lag-CRP curves for "in the zone" vs. "out of the zone" attentional states based on successive recall of items encoded in the same state (Fig. 7B; see Fig. 7A and Table 2 for overall lag-CRP). We conducted a three-way repeated-measures ANOVA with attentional state ("in the zone" vs. "out of the zone"), absolute lag (1 to 29), and direction (forward vs. backward) as factors (results summarized in Table 3).

We found a significant main effect of absolute lag ($F_{8.97,\ 600.97} = 16.22$, $p < 0.0001$, $\eta_p^2 = 0.20$, BF < 1/150): during recall, individuals were more likely to transition to items that were encoded nearby vs. farther away. We also found a significant main effect of direction ($F_{1,67} = 6.24$, $p = 0.02$, $\eta_p^2 = 0.09$, BF = 1.96) and an interaction between direction and absolute lag ($F_{12.47,\ 835.32} = 2.49$, $p = 0.003$, $\eta_p^2 = 0.04$, BF = 0.56). Thus, participants were more likely to recall items in the forward vs. backward direction, with this asymmetry being more pronounced for closer vs. farther lags.

We next examined main effects and interactions involving attentional state. There was no main effect of attentional state ($F_{1,67} = 0.32$, $p = 0.57$, $\eta_p^2 = 0.005$, BF = 31.22). There was also no interaction between attentional state and direction ($F_{1,67} = 0.08$, $p = 0.79$, $\eta_p^2 = 0.001$, BF = 26.88), no interaction between attentional state and absolute lag ($F_{12.37,\ 829.07} = 1.06$, $p = 0.39$, $\eta_p^2 = 0.02$, BF > 150), nor a three-way interaction between absolute lag, direction, and attentional state ($F_{13.11,\ 878.44} = 1.33$, $p = 0.19$, $\eta_p^2 = 0.02$, BF > 150). Hence, we did not see any differences in temporal contiguity or forward asymmetry bias of recall for

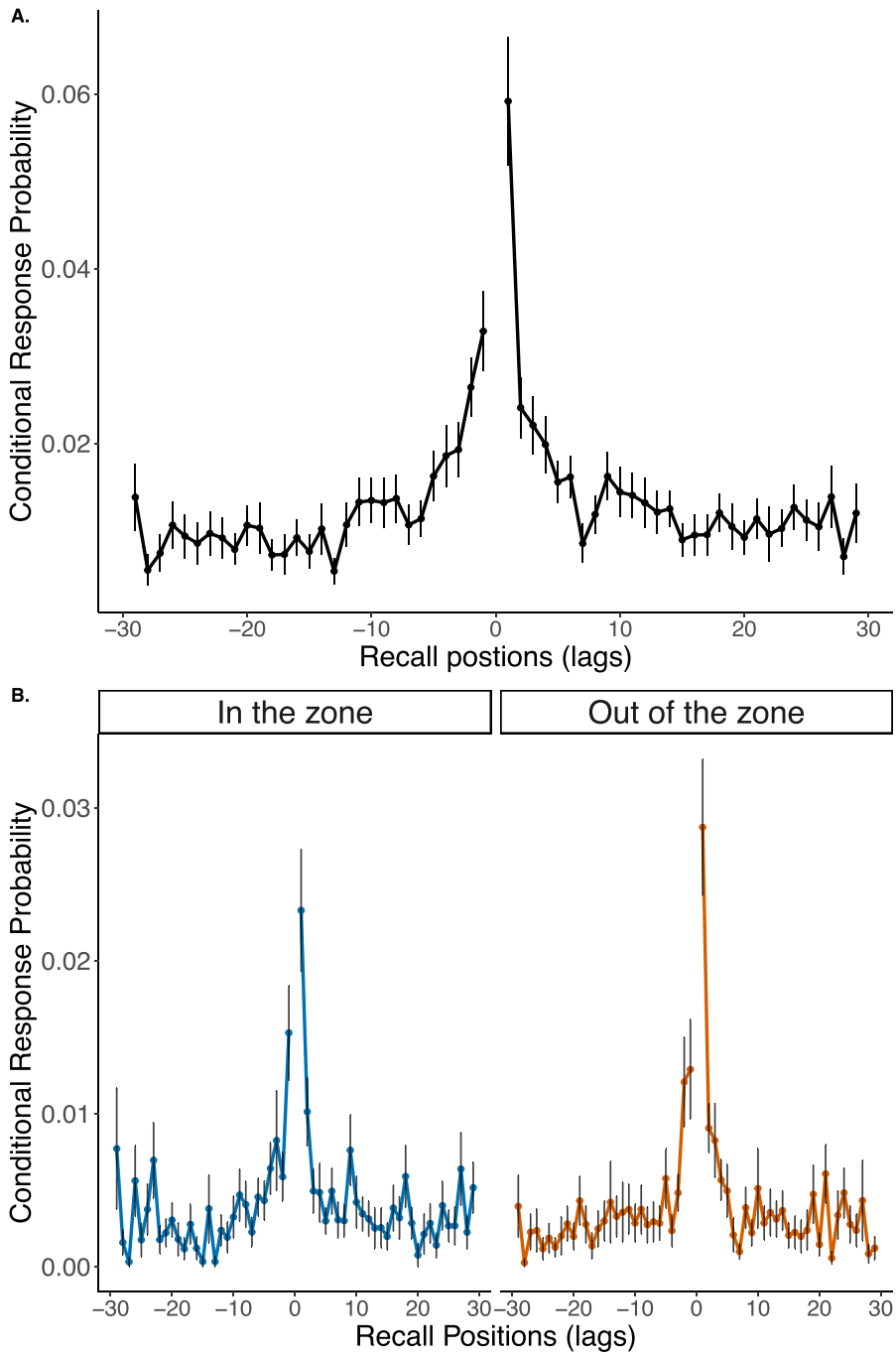

Fig. 7. Lag-CRP curves overall and by attentional state for Study 2. **A.** Overall lag-CRP curve across participants and blocks. **B.** Lag-CRP curves plotted separately for items encoded "in the zone" (left) and "out of the zone" (right). There was no difference between the two attentional states in the temporal organization of recall (neither temporal contiguity nor forward asymmetry). Error bars represent the standard error.

items encoded "in the zone" vs. "out of the zone".

We conducted a follow-up analysis to compare the two attentional states at the lags of ±1. This was done to examine whether including all lags in our repeated-measures ANOVA masked differences between the states that were more specific to close recall transitions. From a two-way repeated-measures ANOVA with lag (+1 vs. -1) and attentional state ("in the zone" vs. "out of the zone") as factors, we found only a significant main effect of lag ($F_{1,67} = 12.09$, $p = 0.0009$, $\eta_p^2 = 0.15$, BF = 0.024). The main effect of attentional state ($F_{1,67} = 0.20$, $p = 0.66$, $\eta_p^2 = 0.003$, BF = 6.89), and the interaction between attentional state and lag, were not statistically significant ($F_{1,67} = 1.28$, $p = 0.26$, $\eta_p^2 = 0.02$, BF = 2.86). The significant main effect of lag (+1 vs. -1) indicated that individuals are more likely to make forward vs. backward transitions at the closest lag. However, this forward asymmetry at the ±1 lags was not different between the two states.

As for Study 1, we conducted a Bayesian model comparison in which we compared the three-way ANOVA reported above to a null model without an attentional state variable. As before, this analysis revealed very strong evidence in favor of the null model (BF > 150).

All together, across these analyses, we replicated the finding that recall is temporally organized. However, this temporal organization was not different between the two attentional states.

3.3.5. No differences in event transition types between the two attentional states

As in Study 1, we next examined recall as a function of the type of transition (Fig. 2 and Section 2.1.2.5). We performed a two-way repeated-measures ANOVA on the conditional probability of recall transitions with transition type (3 levels) and attentional state ("in the zone" vs. "out of the zone") as factors. We found a main effect of transition type ($F_{1.04,\ 69.97} = 46.76$, $p < 0.0001$, $\eta_p^2 = 0.41$, BF < 1/150). The main effect of attentional state ($F_{1,67} = 0.37$, $p = 0.54$, $\eta_p^2 = 0.006$, BF = 7.77) and the interaction between attentional state and transition type ($F_{1.05,\ 70.43} = 0.41$, $p = 0.53$, $\eta_p^2 = 0.006$, BF = 15.08) were not statistically significant. Thus, each type of recall transition was not different for items encoded "in the zone" and "out of the zone" (Fig. 8).

Given the main effect of transition type, we conducted follow-up *t*-tests to compare them, collapsing across the two attentional states, as we did for Study 1. We replicated Study 1 in observing that "same state, same event" transitions were significantly more likely than "same state, different event" transitions ($t_{67} = 7.03$, $p < 0.0001$, Cohen's dz. = 0.85, 95% CI [0.007, 0.013]) and "different state, different event" transitions ($t_{67} = 6.72$, $p < 0.0001$, Cohen's dz. = 0.82, 95% CI [0.007, 0.012]). "Different state, different event" transitions were also significantly more likely than "same state, different event" transitions ($t_{67} = 2.34$, $p = 0.02$, Cohen's dz. = 0.28, 95% CI [0.00008, 0.001]). This pattern is consistent with temporally organized recall, in that "same state, same event" items are the closest to one another, while "same state, different event" items are necessarily never in contiguous events (Fig. 2). As for Study 1, we conducted a Bayesian model comparison in which we compared the two-way ANOVA reported above to a null model without an attentional state variable. This analysis revealed strong evidence in favor of the null model (BF = 109.28).

Thus, across these event transition analyses, we failed to find evidence for our hypotheses regarding differences between "in the zone" and "out of the zone" states.

3.3.6. Secondary analysis: faster vs. slower RTs

As for Study 1, we examined task performance and subsequent memory with a secondary analysis in which encoding RTs were divided, by a median split, into those that were faster vs. those that were slower (see Section 2: **Study 1** for procedural details). This allowed us to determine if this alternative characterization of attentional states reveals differences in recall organization that were masked by our VTC analysis, which considers RTs that are too fast or too slow to be a poor attentional state.

There was no difference in online errors during attentional states characterized by faster vs. slower RTs ($t_{67} = 0.96$, $p = 0.34$, Cohen's dz. = 0.12, 95% CI [−0.40, 1.13], BF = 4.84) and recall was also not different between these two attentional states ($t_{67} = 1.41$, $p = 0.16$, Cohen's dz. = 0.18, 95% CI [−0.35, 2.07], BF = 2.92).

The lag-CRP analysis revealed only a main effect of absolute lag ($F_{11.94,\ 799.95} = 9.60$, $p < 0.0001$, $\eta_p^2 = 0.13$, BF < 1/150); all other main effects and interactions were not statistically significant (all ps > 0.088; all other BFs > 10.53).

The event segment analysis revealed a main effect of transition type ($F_{1.05,\ 70.17} = 46.80$, $p < 0.0001$, $\eta_p^2 = 0.41$, BF < 1/150), reflecting significantly more "same state, same event" transitions compared to "same state, different event" transitions ($t_{67} = 6.91$, $p < 0.0001$, Cohen's dz. = 0.84, 95% CI [0.007, 0.0134]) and "different state, different event" transitions ($t_{67} = 6.88$, $p < 0.0001$, Cohen's dz. = 0.83, 95% CI [0.007, 0.0127]). There was no significant difference between "different state, different event" and "same state, different event" transitions ($t_{67} = 1.83$, $p = 0.07$, Cohen's dz. = 0.22, 95% CI [−0.00004, 0.001]). There was no main effect of attentional state ($F_{1,67} = 1.59$, $p = 0.21$, $\eta_p^2 = 0.023$, BF = 4.45) nor an attentional state by transition type interaction ($F_{1.03,68.96} = 1.62$, $p = 0.21$, $\eta_p^2 = 0.24$, BF = 4.90).

As for Study 1, we performed a Bayesian model comparison for the lag-CRP analysis. We compared the model reported above to a null model without an attentional state variable. This comparison revealed very strong evidence in favor of the null model (BF > 150). The same approach for the event segment analysis revealed strong evidence in favor of the null model (BF = 22.13).

Thus, like the VTC analysis, this alternative characterization of attentional states also failed to reveal the predicted differences in the temporal organization of recall. However, unlike the VTC analysis, this approach was not successful in detecting differences in online performance based on attentional state. Thus, this alternative approach may not be as powerful in identifying fluctuating attentional states; the null effects on memory are therefore more difficult to interpret.

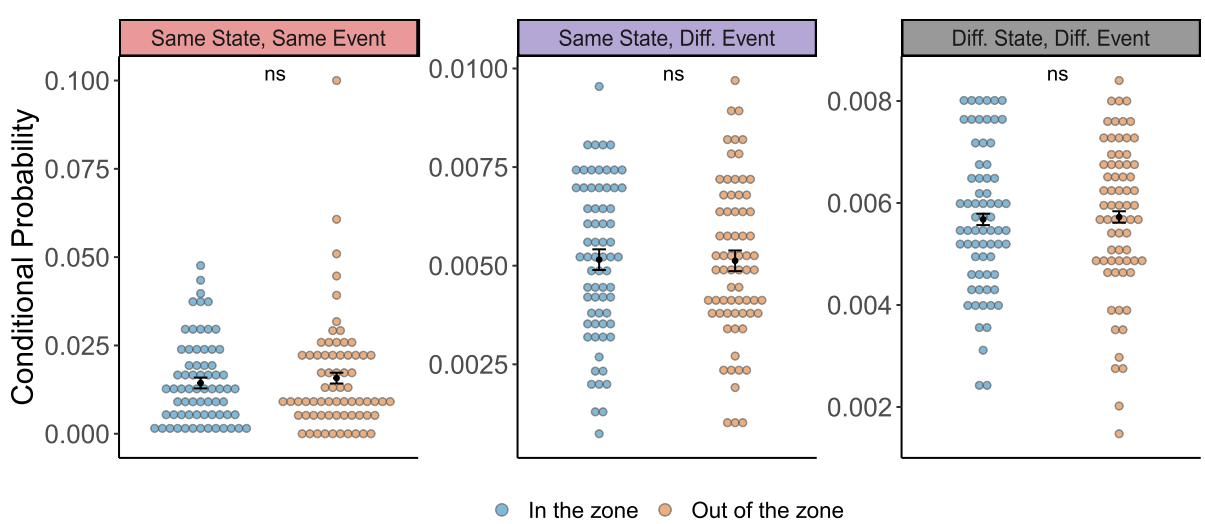

Fig. 8. Recall transitions as a function of event type in Study 2. Recall transitions are shown based on whether they occurred within an "event segment' of a particular attentional state (same state, same event), across event segments of a given attentional state (same state, different event), or between attentional states (different state, different event; see Fig. 2). There were no significant differences between the two attentional states in any transition type. Individual points indicate the conditional probability of each transition type (i.e., the number of times each transition type occurred divided by the number of opportunities to make a transition of that type) for each individual, separately for items encoded "in the zone" and "out of the zone". Black dots indicate the mean; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". ns = not statistically significant.

3.4. Discussion

In Study 2, we made adjustments to our task to try to encourage stronger attentional fluctuations, in particular stronger "zoning out". As before, we hypothesized that "in the zone" attentional states, vs. "out of the zone" states, aid in maintenance of temporal context representations, thus facilitating temporally organized recall. However, we did not find any evidence to support this hypothesis. While we observed differences in encoding task performance, with more errors for "out of the zone" states (replicating prior work by Esterman et al., 2013), we did not find any differences in our recall measures of interest. There was no statistically significant difference in overall recall performance for items encoded during "in the zone" vs. "out of the zone" attentional states. Furthermore, while we replicated prior work in showing both temporal contiguity and forward asymmetry effects in recall, we did not see any differences between the two attentional states in these effects. Finally, we found no differences between the two attentional states in recall transitions between different event types (Fig. 2).

Taken together, we failed to find any evidence that "in the zone" vs. "out of the zone" attentional states have a differential impact on the temporal organization of recall. This was despite our changes to study design that made it more similar to the original gradCPT procedure (see Section 2.3: Study 1 - Discussion).

Why did we fail to find effects of attentional states on recall organization in Study 2? One possibility is that, despite longer blocks, making semantic judgements (food vs. non-food item) may have been challenging enough to engage participants' sustained attention and hence, did not produce strong periods of "zoning out" as we hoped. In contrast, the traditional gradCPT often has participants perform a more perceptual task (e.g., male or female face). The semantic judgement may also encourage semantic clustering (Long & Kahana, 2017), which may interfere with our ability to detect differences in the temporal structure of recall between the two attentional states. Finally, it is possible that the gradual transitions between items — a core part of the gradCPT — have an unintended effect of making the task somewhat engaging, in that participants can try to identify objects at lower and lower opacities as the task goes on. Such gradual transitions also make our task less similar to standard list-learning recall tasks, which present each item in isolation. We address these limitations in Study 3.

4. Study 3

4.1. Overview

In Study 3, we made the following changes to address the limitations above. First, we changed the encoding judgement to be a perceptual one (i.e., is the image color or grayscale?) instead of a semantic one. This aligned our approach with the traditional gradCPT, which uses perceptual judgements (e.g., male or female face?). This judgement is also relatively simple and unambiguous, reducing the likelihood that RT variability may be related to task difficulty. We also made this change to ensure that the task judgement does not encourage semantic clustering of items (Long & Kahana, 2017). For example, if participants in Study 2 attempted to cluster food items together, and non-food items together, we may have had less of an opportunity to observe subtle differences in temporal clustering between the two attentional states. Second, we removed the gradual transitions between items, and replaced this transition with a relatively long presentation duration (3 s on the screen, with a 2 s inter-item interval). This change makes our design more similar to standard memory tasks, and may additionally make the task less engaging. Finally, in an attempt to improve recall performance, we reduced the length of our study phase (60 items instead of the 80 in Study 2).

4.2. Methods

4.2.1. Design

4.2.1.1. Participants. We report data from 68 participants ($M_{age} = 20.09 \pm 2.20$, $M_{education} = 14.03 \pm 1.46$; see Table 1 for demographics). We do not report data from an additional 4 participants, who were excluded due to low response rate during the encoding task (<80%, $N = 1$), age >40 years (N = 1), recall recording issues (N = 1) and incomplete participation (N = 1). All participants were recruited from the Columbia University participant pool and participated in an online version of the task hosted on the Gorilla platform (www.gorilla.sc; Anwyl-Irvine et al., 2020). Informed consent was obtained in accordance with the Columbia University Institutional Review Board.

4.2.1.2. Stimuli. Stimuli were identical to Study 2 with the following exceptions. We chose 180 objects from the pre-curated databases. Of these, 10% of the images (18 images) were converted to grayscale. 90% of the images (162 images) were in color. The 180 images were divided into 3 lists of 60 images each (6 grayscale images, 54 color images). This was done manually by ensuring there were an equal number of items from a category (e.g., tools, furniture) in each list.

4.2.1.3. Procedure. The procedure was identical to Study 2 with the following exceptions. In each study phase, participants viewed 60 trial-unique items from the created lists. Each image was presented for 3 s followed by a fixation cross during the 2 s inter-item interval (i.e., there was no gradual fading between images). Participants were asked to judge if each image was in color or grayscale: They pressed a button when it was in color, but withheld their response when it was in grayscale. Participants therefore habitually pressed one response 90% of the time. The distractor and recall phases were identical to the ones in Study 2 (Fig. 1).

4.2.2. Analyses

All analyses were identical to Study 2, except that response times (RTs) were defined from the onset of the static image on each trial without any assignment algorithm (because there was no fading in).

4.3. Results

4.3.1. Defining attentional states at encoding

In the encoding task, participants viewed images and judged each as being in color (with a button press) or grayscale (by withholding their response). Overall, mean response time (RT; defined from image onset) was 1.33 s (SD = 0.20). Median RT was 1.30s. As in Studies 1 and 2, we performed a variance time course analysis on the encoding phase RTs. Fig. 9A shows the VTC analysis for one sample participant in Study 3.

The mean length of an "in the zone" segment was 3.85 trials (SD = 0.53) and the mean length of an "out of the zone" segment was 3.81 trials (SD = 0.47; Note that each trial was 5 s long). The mean number of fluctuations within a block (i.e., the number of times participants transitioned from one state to another) was 14.84 (SD = 2.10). The number of trials within a segment ranged from 1 to 18 for "in the zone" states and from 1 to 20 for "out of the zone" states, across all blocks and participants.

As for prior Studies, we tested whether RT variability changed over blocks. Unlike Studies 1 and 2, however, RT variability was not significantly different across blocks ($F_{2,134} = 0.24$, $p = 0.79$, $\eta_p^2 = 0.004$). This suggests that RT variability, and the VTC approach, may not be as meaningful in this Study, perhaps because we removed the gradual image transitions that are important in the gradCPT. To further test if this is the case, we examined error rates for "in the zone" and "out of the zone" states, below.

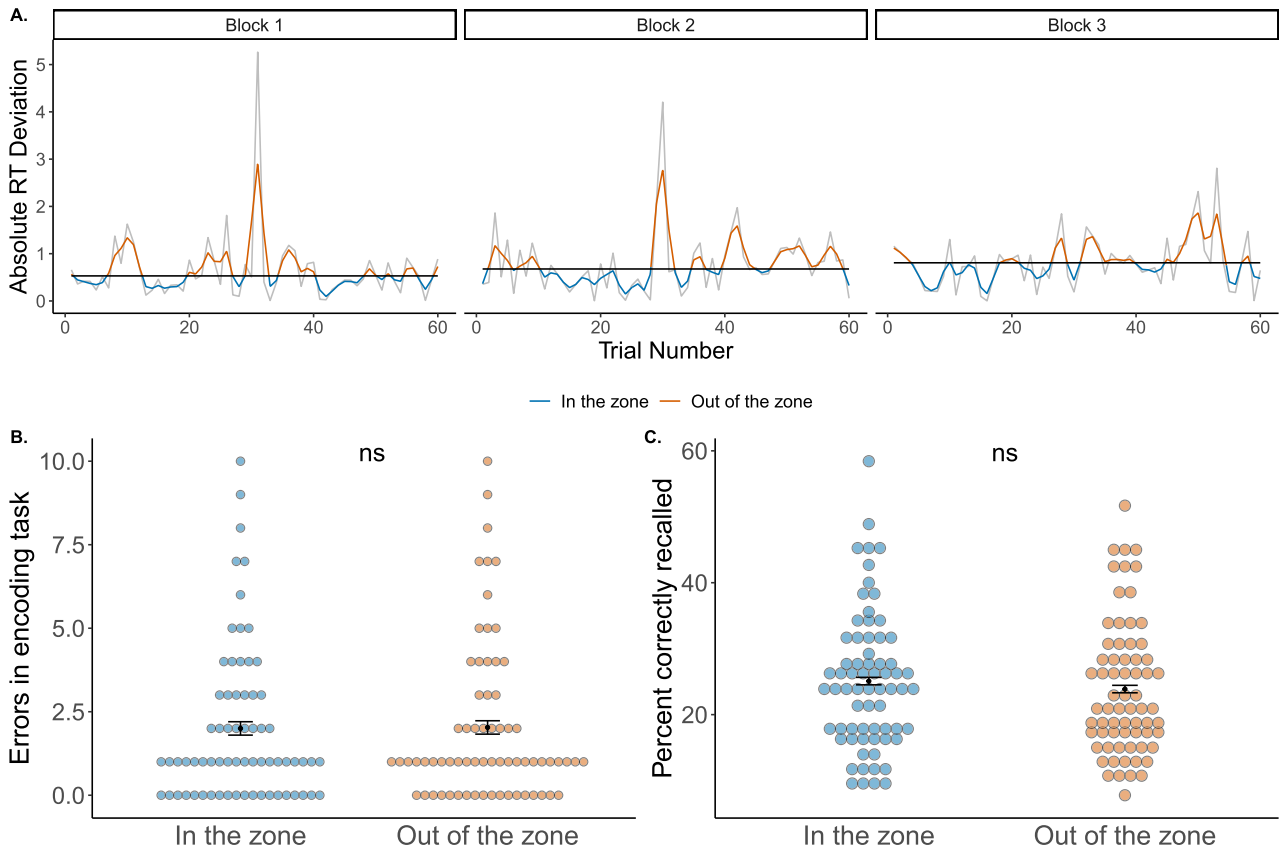

Fig. 9. Encoding task performance and recall performance in Study 3. Neither encoding errors nor recall performance differ between attentional states. **A.** Variance Time Course (VTC) analysis for a sample participant depicting "in the zone" (blue) and "out of the zone" (orange) attentional states. Horizontal black lines indicate the median absolute RT deviation per block. Gray curves indicate raw (unsmoothed) RT deviation per block. **B.** Individual points show the number of encoding judgement errors made by each participant during "in the zone" and "out of the zone" attentional states. The number of encoding errors did not differ between "in the zone" vs. "out of the zone" attentional states. **C.** Individual points show the percentage of items correctly recalled by each participant as a function of whether items were encoded "in the zone" or "out of the zone". There was no difference in recall performance between the two states. Black points in panels B & C indicate the mean of the measure; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". ns = not statistically significant. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.3.2. No difference in encoding errors during "in the zone" vs "out of the zone" attentional states

Participants performed very well on the encoding task ("Is this image in color or grayscale?"). They responded to 98.80% (SD = 2.58%; Median = 100%) of the "go" color image trials, which required a response. Mean accuracy (defined as correct responses on "go" trials and withheld responses on "no-go" trials) was 97.45% (SD = 2.61%; Median = 98.34%).

We next examined errors in the encoding task as a function of attentional state. Surprisingly, we failed to replicate Studies 1 and 2: A paired samples *t*-test revealed that the number of encoding errors during an "out of the zone" attentional state (mean ± SD: 2.03 ± 2.40) was not significantly different from that in an "in the zone" attentional state (2.00 ± 2.34; $t_{67} = 0.10$, $p = 0.92$, Cohen's dz. = 0.012, 95% CI [−0.54,0.60], BF = 7.47, Fig. 9B). Thus, unlike Studies 1 and 2, and unlike prior studies (Rosenberg et al., 2013, Esterman et al., 2013), the VTC analysis was unsuccessful in identifying fluctuations between better and worse attentional states.

As in prior Studies, we next examined whether there was post-error slowing or speeding. Unlike Studies 1 and 2, we found no difference in RTs before vs. after an error (omission errors: $\beta = 42.24$, S.E. = 78.70, $t_{387.25} = 0.54$, $p = 0.59$, 95% CI [−112.00, 196.49]; commission errors: $\beta = -57.49$, S.E. = 83.22, $t_{138.99} = -0.69$, $p = 0.49$, 95% CI [−220.61, 105.63]). This offers further evidence that RTs in this study may not reliably index attentional fluctuations.

Thus, although removing gradual transitions made our design more similar to standard memory tasks, the abrupt image onsets may have captured attention, reducing "zoning out" (Esterman et al., 2013). We return to this issue in the **Discussion (Section 4.4)**. Despite this null result, we report the rest of Study 3 results for completeness.

4.3.3. Overall recall does not differ between the two attentional states

We next examined recall performance. Mean recall (i.e., the percentage of items correctly recalled across all blocks) was 24.32% (SD = 9.78).

We then separately examined recall performance based on whether items were encoded "in the zone" or "out of the zone". As in Studies 1 and 2, we did not find a significant difference between "in the zone" (Mean ± SD: 25.09% ± 10.40) and "out of the zone" (23.88% ± 10.10) attentional states ($t_{67} = 1.51$, $p = 0.14$, Cohen's dz. = 0.18, 95% CI [−0.39, 2.81], BF = 2.56, Fig. 9C). Thus, attentional states in this Study did not differ in either online task performance or in subsequent recall.

4.3.4. No differences in temporal contiguity or forward asymmetry between the two attentional states

As in Studies 1 and 2, we examined lag-CRP curves to explore the temporal organization of recall. This allowed us to determine whether the structure of memory differed between the two attentional states, even if overall memory performance did not.

As before, we constructed separate lag-CRP curves for "in the zone"

vs. "out of the zone" attentional states based on successive recall of items encoded in the same state (Fig. 10B; see Fig. 10A and Table 2 for overall lag-CRP). We conducted a three-way repeated-measures ANOVA with attentional state ("in the zone" vs. "out of the zone"), absolute lag (1 to 29), and direction (forward vs. backward) as factors (results summarized in Table 3).

We found a significant main effect of absolute lag ($F_{10.97,735.01} = 9.98$, $p < 0.0001$, $\eta_p^2 = 0.13$, BF < 1/150): during recall, individuals were more likely to transition to items that were encoded nearby vs. farther away. There was no main effect of direction ($F_{1,67} = 0.066$, $p = 0.80$, $\eta_p^2 = 0.001$, BF = 38.88), but there was a significant interaction between direction and absolute lag ($F_{14.01,938.37} = 2.17$, $p = 0.007$, $\eta_p^2 = 0.03$, BF = 18.22): participants were more likely to recall items in the forward vs. backward direction, with this asymmetry being larger for closer vs. farther lags. There was no main effect of attentional state ($F_{1,67} = 1.68$, $p = 0.20$, $\eta_p^2 = 0.025$, BF = 12.39). There was also no interaction between attentional state and direction ($F_{1,67} = 0.12$, $p = 0.73$, $\eta_p^2 = 0.002$, BF = 26.99), no interaction between attentional state and absolute lag ($F_{14.74,987.61} = 0.82$, $p = 0.66$, $\eta_p^2 = 0.013$, BF > 150), nor a three-way interaction between absolute lag, direction, and attentional state ($F_{15.20,1018.34} = 0.95$, $p = 0.51$, $\eta_p^2 = 0.014$, BF > 150). Hence, we did not see any differences in recall organization — neither temporal contiguity nor forward asymmetry bias — based on attentional state at encoding.

As before, we conducted a follow-up analysis to examine differences

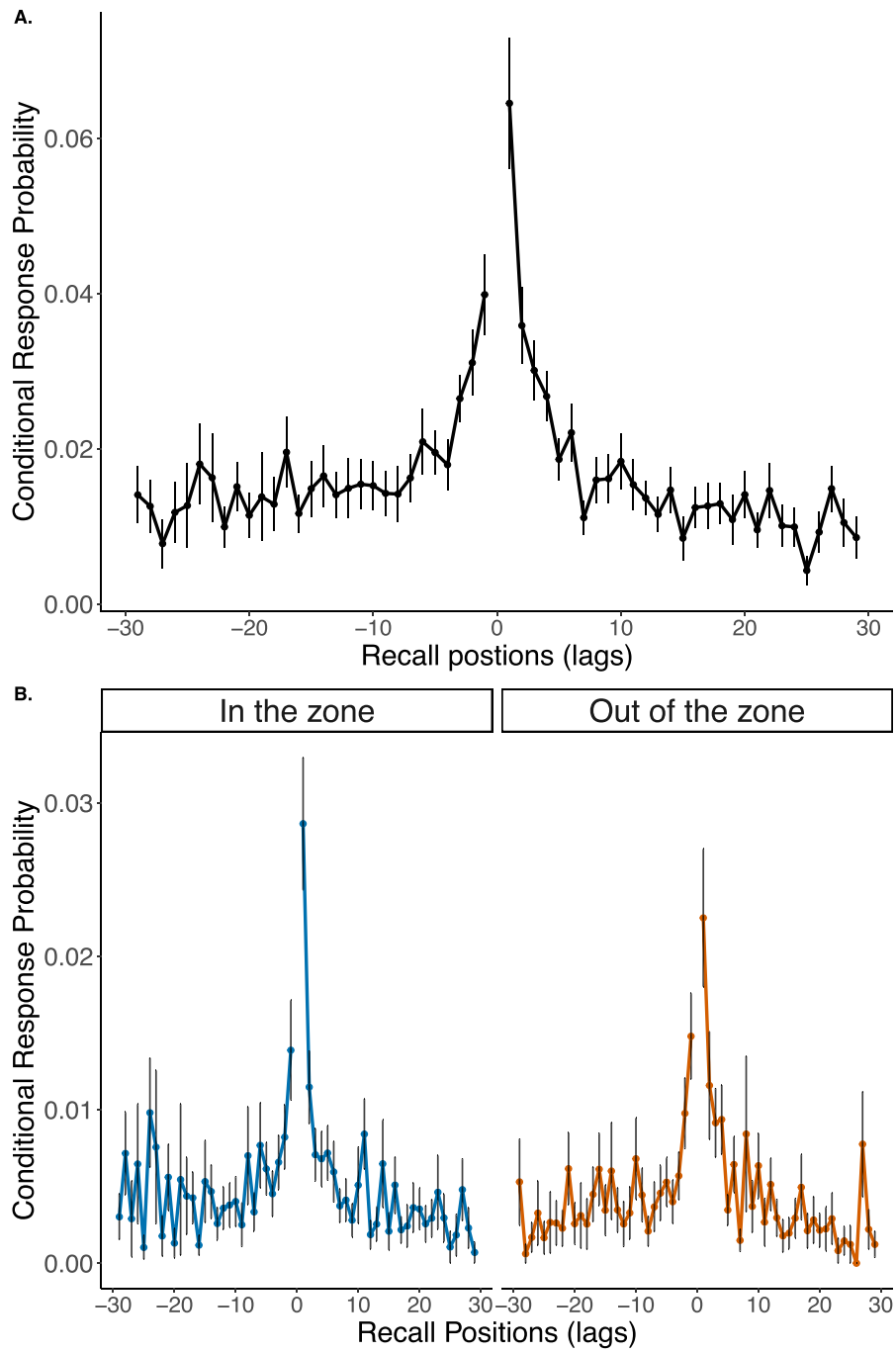

Fig. 10. Lag-CRP curves overall and by attentional state for Study 3. **A.** Overall lag-CRP curve across participants and blocks. **B.** Lag-CRP curves plotted separately for items encoded "in the zone" (left) and "out of the zone" (right). There was no difference between the two attentional states in the temporal organization of recall (neither temporal contiguity nor forward asymmetry). Error bars represent the standard error.

between the two attentional states at the nearby lags of ±1. From a two-way repeated-measures ANOVA with lag (+1 vs. -1) and attentional state ("in the zone" vs. "out of the zone") as factors, we found only a significant main effect of lag ($F_{1,67} = 12.57$, $p = 0.0007$, $\eta_p^2 = 0.16$, BF = 0.028). The main effect of attentional state ($F_{1,67} = 0.55$, $p = 0.46$, $\eta_p^2 = 0.008$, BF = 6.36), and the interaction between attentional state and lag, was not statistically significant ($F_{1,67} = 1.32$, $p = 0.26$, $\eta_p^2 = 0.019$, BF = 3.08). The significant main effect of lag (+1 vs. -1) confirms that individuals are more likely to make forward vs. backward transitions at the closest lag. However, forward asymmetry at the ±1 lags was not different between the two states.

As for prior Studies, we conducted a Bayesian model comparison in which we compared the three-way ANOVA reported above to a null model without an attentional state variable. Once again, this analysis revealed very strong evidence in favor of the null model (BF > 150).

We therefore once again replicated the finding that recall is temporally organized. However, the temporal organization of recall was not different between the two attentional states. This replicates the null findings from the lag-CRP analyses in Studies 1 and 2.

4.3.5. No differences in event transition types between the two attentional states

As in Studies 1 and 2, we next examined recall transitions as a function of the type of event segment (see Fig. 2 and Section 2.1.2.5).

To do this, we performed a two-way repeated-measures ANOVA with transition type (3 levels) and attentional state ("in the zone" vs. "out of the zone") as factors. As before, we hypothesized that "same state, same event" and "same state, different event" transitions may be more likely for items encoded "in the zone" vs. "out of the zone".

We found a main effect of transition type ($F_{1.06,70.93} = 34.19$, $p < 0.0001$, $\eta_p^2 = 0.34$, BF < 1/150). The main effect of attentional state ($F_{1,67} = 2.30$, $p = 0.13$, $\eta_p^2 = 0.03$, BF = 4.41), and the interaction between attentional state and transition type ($F_{1.14,76.32} = 1.32$, p = 0.26, $\eta_p^2 = 0.02$, BF = 10.66), were not statistically significant. Thus, each type of recall transition was not differentially likely for items encoded "in the zone" and "out of the zone" (Fig. 11).

As for prior Studies, we conducted follow-up *t*-tests to understand the main effect of transition types, collapsing across the two attentional states. We found that: 1) "same state, same event" transitions were significantly more likely compared to "same state, different event" transitions ($t_{67} = 6.10$, $p < 0.0001$, Cohen's dz. = 0.74, 95% CI [0.008, 0.015]) and "different state, different event" transitions ($t_{67} = 5.65$, p < 0.0001, Cohen's dz. = 0.69, 95% CI [0.006, 0.013]), and 2) "different state, different event" transitions were significantly more likely compared to "same state, different event" transitions ($t_{67} = 3.47$, $p = 0.0009$, Cohen's dz. = 0.42, 95% CI [0.0005, 0.002]).

As for prior Studies we conducted a Bayesian model comparison in which we compared the two-way ANOVA reported above to a null model without an attentional state variable. This analysis again revealed strong evidence in favor of the null model (BF = 44.01). As in prior Studies, therefore, we did not find any evidence that recall transition types differed between the two attentional states.

4.3.6. Secondary analysis: faster vs. slower RTs

As before, we examined task performance and subsequent memory with a secondary analysis. Rather than using the VTC analysis to characterize "in the zone" and "out of the zone" states, we divided encoding RTs, by a median split, into those that were faster vs. those that were slower. We found no difference in error rates during attentional states characterized by faster vs. slower RTs ($t_{67} = 0.85$, $p = 0.40$, Cohen's dz. = 0.10, 95% CI [−0.32, 0.79], BF = 5.30). This is consistent with Studies 1 and 2, in which this analysis also failed to reliably predict online errors, and also consistent with the VTC analysis above, which likewise failed to find differences in error rates for "in the zone" vs. "out of the zone" states in this Study.

This analysis did, however, reveal a difference in subsequent recall performance: attentional states associated with slower (vs. faster) RTs at encoding were associated with better subsequent recall ($t_{67} = 2.18$, $p = 0.033$, Cohen's dz. = 0.26, 95% CI [0.14, 3.13], BF = 0.82). The lag-CRP analysis nevertheless failed to reveal differences in the temporal organization of recall: there was only a main effect of absolute lag ($F_{11.90, 797.25} = 9.23$, p < 0.0001, $\eta_p^2 = 0.12$, BF < 1/150; all other ps > 0.06; all other BFs > 24.81).

The event segment analysis revealed a main effect of transition type ($F_{1.04, 69.80} = 29.90$, p < 0.0001, $\eta_p^2 = 0.31$, BF < 1/150), reflecting significantly more "same state, same event" transitions compared to "same state, different event" transitions ($t_{67} = 5.55$, p < 0.0001, Cohen's dz. = 0.67, 95% CI [0.007, 0.015]) and "different state, different event"

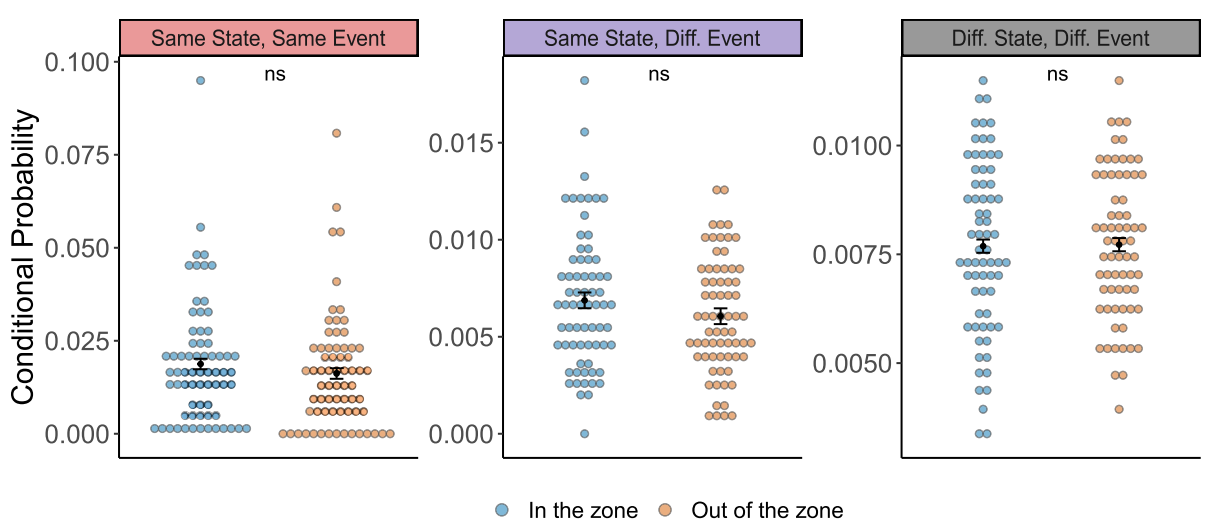

Fig. 11. Recall transitions as a function of event type in Study 3. Recall transitions are shown based on whether they occurred within an "event segment' of a particular attentional state (same state, same event), across event segments of a given attentional state (same state, different event), or between attentional states (different state, different event; see Fig. 2**)**. There were no significant differences in any transition type between the two attentional states. Individual points indicate the conditional probability of each transition type (i.e., the number of times each transition type occurred divided by the number of opportunities to make a transition of that type) for each individual, separately for items encoded "in the zone" and "out of the zone". Black dots indicate the mean; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". ns = not statistically significant.

transitions ($t_{67} = 5.46$, $p < 0.0001$, Cohen's dz. = 0.66, 95% CI [0.007, 0.015]). There was no significant difference between "different state, different event" and "same state, different event" transitions ($t_{67} = 1.50$, $p = 0.14$, Cohen's dz. = 0.18, 95% CI [−0.0002, 0.001]). There was no main effect of attentional state ($F_{1,67} = 2.55$, $p = 0.11$, $\eta_p^2 = 0.04$, BF = 3.60) nor an attentional state by transition type interaction ($F_{1.12,74.90} = 1.55$, $p = 0.22$, $\eta_p^2 = 0.02$, BF = 8.60).

As for Studies 1 and 2, we performed a Bayesian model comparison for the lag-CRP analysis, in which we compared the model reported above to a null model without an attentional state variable. This comparison revealed very strong evidence in favor of the null model (BF > 150). The same approach for the event segment analysis revealed strong evidence in favor of the null model (BF = 31.77).

Thus, like the VTC analysis, this alternative characterization of attentional states also failed to reveal statistically robust differences in the temporal organization of recall.

4.4. Discussion

In Study 3, we made further adjustments to our task to encourage more "zoning out" and to make our procedure more similar to standard list-learning recall tasks. We again tested the hypothesis that "in the zone", vs. "out of the zone", states aid in maintenance of temporal context representations, thus encouraging temporally organized recall. However, we did not find any evidence to support this hypothesis.

Contrary to Studies 1 and 2, we did not observe any differences in encoding task performance across the two attentional states. Why might this be the case? First, we removed the gradual transitions between images. Although this change made our task more similar to standard list-learning recall tasks, the abrupt image onsets could have captured attention, thus preventing participants from "zoning out" (Esterman et al., 2013; Rosenberg et al., 2013). Second, long and variable interstimulus durations result in more attentional lapses than short, fixed durations (Unsworth, Robison, & Miller, 2018). Our short and fixed ITI (2 s) may therefore have made it less likely that participants would zone out. Finally, we changed the encoding judgement to a simple perceptual judgement. Although this change made our task more similar to the traditional gradCPT, which often uses perceptual judgements, it had the effect of increasing the accuracy of encoding judgements. Having such a low error rate (~2 errors, on average, per participant in each attentional state) may have hurt our chances of seeing differences between the two states. This low error rate may have been due to the saliency of color changes, which might capture attention even if a participant had previously been zoning out. The lack of a difference in errors between "in the zone" and "out of the zone" states makes it difficult to interpret the lack of differences between them in subsequent recall organization: we may not have adequately captured attentional fluctuations based on RT. This prompted us to conduct Study 4, which is similar to Study 2, with some minor changes to once again try to induce stronger attentional fluctuations.

5. Study 4

5.1. Overview

In Study 3, online task performance was not different between "in the zone" and "out of the zone" states. This makes the null effects in recall organization difficult to interpret, because we may not have successfully differentiated between better and worse attentional states. We therefore conducted another study with the gradual transitions used in Studies 1 and 2, in which we successfully replicated prior work showing more errors during "out of the zone" states (Esterman et al., 2013; Rosenberg et al., 2013). Study 4 used a similar design as Study 2 (go/no-go procedure, with responses for non-food items and withheld responses for food items). Two changes were made: first, there were 2 blocks of 120 images each (instead of 3 blocks of 80 images each), and second, the trial duration (from the onset of an image fading in until it became 100% clear) was reduced from 6 s to 4 s. Both of these changes were implemented to bring our design closer to the traditional gradCPT, which typically uses fast presentation durations and many trials. In this way, we hoped to induce stronger attentional fluctuations. The increase in block length, however, means that our procedure deviates even more from standard list-learning recall studies, which typically use short lists (Murdock Jr., 1962; Cortis, Dent, Kennett, & Ward, 2015; also see Healey et al., 2019). We return to this point in the Section 6: **General Discussion**.

5.2. Methods

5.2.1. Design

5.2.1.1. Participants. Pilot data using the Study 4 procedure revealed that participants had worse recall performance than our earlier studies (likely because Study 4 blocks were longer than those in our prior studies). We therefore opted to collect a larger sample size, so that summed recall performance across all participants would be comparable to Study 2 (Baker et al., 2020). We report data from 124 participants ($M_{age} = 21.42 \pm 6.26$, $M_{education} = 13.63 \pm 1.54$; see Table 1 for demographics). We do not report data from an additional 32 participants, who were excluded due to image loading errors ($N = 1$), low response rate during the encoding task (<80%, $N = 20$), recall recording issues ($N = 10$), and no recall (N = 1). Of the final sample, 5 participants were recruited through Prolific (www.prolific.co) and the rest (119 participants) were recruited from the Columbia University participant pool. All participants completed an online version of the task hosted on the Gorilla platform (www.gorilla.sc; Anwyl-Irvine et al., 2020). Informed consent was obtained in accordance with the Columbia University Institutional Review Board.

5.2.1.2. Stimuli. Stimuli were identical to Study 2, except that the 240 images were divided into 2 lists of 120 images each (12 food, 108 non-food images).

5.2.1.3. Procedure. The procedure was identical to Study 2 with the following exceptions. The experiment consisted of 2 blocks, each of which included a study phase, a distractor phase, and a recall phase (Fig. 1). In each study phase, participants viewed 120 trial-unique items, which transitioned slowly from one into another. Trial duration was 4 s instead of 6 s.

The distractor phase was identical to that in Study 2. The recall phase was similar to Study 2, except that participants were given the option to recall for a longer duration. Participants were initially given 4 min (broken into 2 recordings of 2 min each) to verbally recall items from the study phase. After the initial 4 min of recording, participants were given the option of recording for an additional 2 min. This was done because the encoding blocks in Study 4 were longer than those in Study 2; thus, we wanted to give participants more time to recall if they needed it.

5.2.2. Analyses

Analyses were identical to Study 2 (Section 3.2.2.).

5.3. Results

5.3.1. Defining attentional states at encoding

In the encoding task, participants viewed images and judged each as being a non-food item (with a button press) or a food item (by withholding their response). Overall, mean response time (RT; defined from image onset) was 2.53 s (SD = 0.36). Median RT was 2.55 s.

As in Studies 1–3, we performed a variance time course analysis on the encoding phase RTs. Fig. 12A shows the VTC analysis for one sample participant in Study 4.

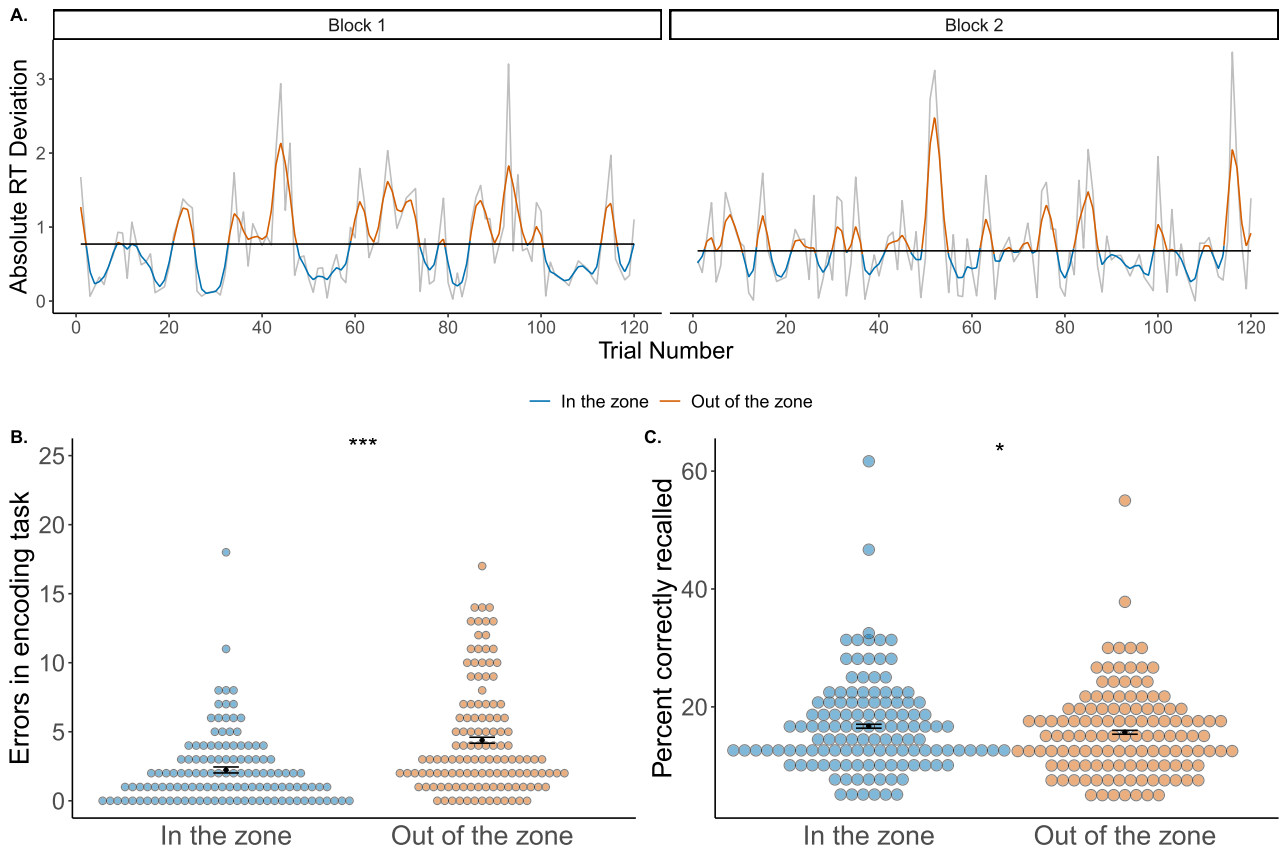

Fig. 12. Encoding task performance and recall performance in Study 4. Both encoding errors and recall performance differ between the two attentional states. **A.** Variance Time Course (VTC) analysis for a sample participant, depicting "in the zone" (blue) and "out of the zone" (orange) attentional states. Horizontal black lines indicate the median absolute RT deviation per block. Gray curves indicate raw (unsmoothed) RT deviation per block. **B.** Individual points show the number of encoding judgement errors made by each participant during "in the zone" and "out of the zone" attentional states. Participants made significantly more encoding errors during the "out of the zone" state. **C.** Individual points show the percentage of items correctly recalled by each participant as a function of whether items were encoded "in the zone" or "out of the zone". There was a small but statistically significant difference in recall performance: recall was higher for items encoded while "in the zone" vs. "out of the zone". Black points in panels B & C indicate the mean of the measure; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". * $p < 0.05$, *** $p < 0.0001$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The mean length of an "in the zone" segment was 4.97 trials (SD = 0.58) and the mean length of an "out of the zone" segment was 4.93 trials (SD = 0.56; Note that each trial was 4 s long). The mean number of fluctuations within a block (i.e., the number of times participants transitioned from one state to another) was 23.45 (SD = 2.67). The number of trials within a segment ranged from 1 to 33 for "in the zone" states and from 1 to 27 for "out of the zone" states, across all blocks and participants.

As for prior Studies, we performed an initial validation check for the VTC approach by testing whether RT variability changed over blocks. As in Studies 1 and 2, RT variability increased across blocks ($t_{123} = 6.79$, $p < 0.0001$, Cohen's dz. = 0.61, 95% CI [68.22, 124.34]), which may be suggestive of increased fatigue or loss of focus as the experiment continued. This supports the use of RT variability to differentiate "in the zone" and "out of the zone" states.

These attentional states were used to examine accuracy on the encoding task and subsequent recall performance, described below.

5.3.2. More encoding errors during "out of the zone" attentional states

Participants once again performed very well on the encoding task ("Is this item a food or a non-food item?"). They responded to 99.16% (SD = 2.10%; Median = 99.54%) of the "go" non-food image trials, which required a response. Mean accuracy (defined as correct responses on "go" trials and withheld responses on "no-go" trials) was 97.08% (SD = 2.55%; Median = 97.92%).

We next examined errors in the encoding task as a function of attentional state. We replicated Studies 1 and 2 and prior studies (e.g., Esterman et al., 2013, 2014; Fortenbaugh et al., 2018; Rosenberg et al., 2013; Rosenberg et al., 2015). Participants made significantly more errors during an "out of the zone" attentional state (mean ± SD: 4.39 ± 4.00) compared to an "in the zone" attentional state (2.23 ± 2.63; $t_{123} = 7.04$, $p < 0.0001$, Cohen's dz. = 0.63, 95% CI [1.55, 2.76], BF < 1/150, Fig. 12B). Thus, the VTC analysis was once again successful in identifying fluctuations between better and worse attentional states.

As in prior Studies, we performed an additional validation check of whether RTs are sensitive to attentional fluctuations in our task, by comparing RTs before vs. after an error. As in Study 2, participants were faster to respond after vs. before an omission error ($\beta = -970.21$, S.E. = 67.82, $t_{650.25} = -14.31$, $p < 0.0001$, 95% CI [−1103.13, −837.29]). Conversely, participants were slower to respond after vs. before a commission error ($\beta = 210.90$, S.E. = 62.80, $t_{461.06} = 3.36$, $p = 0.0009$, 95% CI [87.82, 333.98]). These results are consistent with participants speeding up after realizing they failed to respond to the previous trial, and slowing down after realizing they incorrectly responded when they should not have. Thus, post-error slowing after commission errors and post-error speeding after omission errors show that RTs can index attentional fluctuations in our task.

5.3.3. Recall performance is better for items encoded during "in the zone" states

We next examined recall performance. Mean recall (i.e., the percentage of items correctly recalled across all blocks) was 16.26% (SD = 7.36).

We then separately examined recall performance based on whether items were encoded "in the zone" or "out of the zone". Unlike Studies 1–3, we found a significant difference between the attentional states, such that participants recalled a higher percentage of items that were encoded during an "in the zone" (Mean ± SD: 16.72% ± 8.12) vs. "out of the zone" (15.69% ± 7.51) attentional state ($t_{123} = 2.21$, $p = 0.03$, Cohen's dz. = 0.20, 95% CI [0.11, 1.95], BF = 0.96, Fig. 12C). Thus, attentional states at encoding were associated with a difference in both online task performance and subsequent recall performance, with both being superior for "in the zone" states. This suggests that the VTC analysis was successful in identifying better vs worse attentional states.

5.3.4. No differences in temporal contiguity or forward asymmetry between the two attentional states

As in prior Studies, we examined lag-CRP curves to explore the temporal organization of recall. This allowed us to determine whether the structure of memory differed between the two attentional states.

As before, we constructed separate lag-CRP curves for "in the zone" vs. "out of the zone" attentional states based on successive recall of items

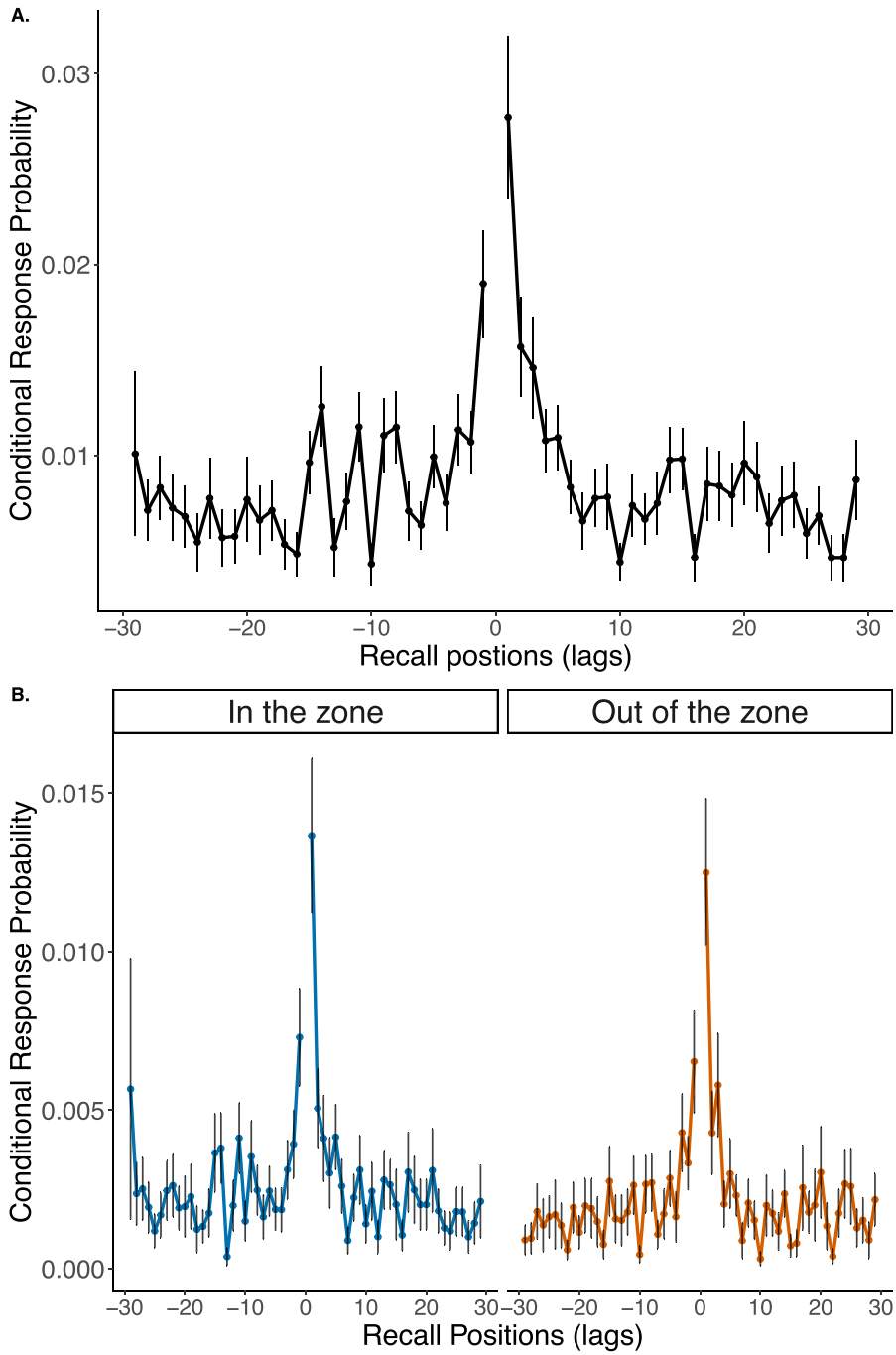

Fig. 13. Lag-CRP curves overall and by attentional state for Study 4. **A.** Overall lag-CRP curve across participants and blocks. **B.** Lag-CRP curves plotted separately for items encoded "in the zone" (left) and "out of the zone" (right). There was no difference between the two attentional states in the temporal organization of recall (neither temporal contiguity nor forward asymmetry). Error bars represent the standard error.

encoded in the same state (Fig. 13B; see Fig. 13A and Table 2 for overall lag-CRP). We conducted a three-way repeated-measures ANOVA with attentional state ("in the zone" vs. "out of the zone"), absolute lag (1 to 29), and direction (forward vs. backward) as factors (results summarized in Table 3).

We found a significant main effect of absolute lag ($F_{10.22,1256.64} = 10.33$, $p < 0.0001$, $\eta_p^2 = 0.08$, BF < 1/150): during recall, individuals were more likely to transition to items that were encoded nearby vs. farther away. There was no main effect of direction ($F_{1,123} = 1.69$, $p = 0.20$, $\eta_p^2 = 0.014$, BF = 24.95). However, there was a significant interaction between direction and absolute lag ($F_{12.59,\ 1549.09} = 2.00$, $p = 0.02$, $\eta_p^2 = 0.016$, BF > 150). This interaction arose because forward asymmetry (forward > backward recall) was more pronounced for closer vs. farther lags.

Unlike Studies 1–3, but consistent with the recall difference between attentional states, there was a main effect of attentional state ("in the zone" > "out of the zone"; $F_{1,123} = 3.97$, $p = 0.0485$, $\eta_p^2 = 0.03$, BF = 2.08). However, as in Studies 1–3, there was no interaction between attentional state and direction ($F_{1,123} = 0.52$, $p = 0.47$, $\eta_p^2 = 0.004$, BF = 29.46), no interaction between attentional state and absolute lag ($F_{13.13,\ 1615.28} = 0.62$, $p = 0.84$, $\eta_p^2 = 0.005$, BF > 150), nor a three-way interaction between absolute lag, direction, and attentional state ($F_{13.16,\ 1618.76} = 0.55$, $p = 0.90$, $\eta_p^2 = 0.0045$, BF > 150). Hence, we did not see any differences in recall organization — neither temporal contiguity nor forward asymmetry bias — based on attentional state at encoding.

As before, we also conducted a follow-up analysis to examine differences between the two attentional states at the nearby lags of ±1. From a two-way repeated-measures ANOVA with lag (+1 vs. -1) and attentional state ("in the zone" vs. "out of the zone") as factors, we found only a significant main effect of lag ($F_{1,123} = 9.69$, $p = 0.002$, $\eta_p^2 = 0.07$, BF = 0.043). The main effect of attentional state ($F_{1,123} = 0.30$, $p = 0.59$, $\eta_p^2 = 0.002$, BF = 8.96), and the interaction between attentional state and lag, was not statistically significant ($F_{1,123} = 0.01$, $p = 0.92$, $\eta_p^2 = 0.00008$, BF = 7.20). The significant main effect of lag (+1 vs. -1) confirms that individuals are more likely to make forward vs. backward transitions at the closest lag. Nevertheless, forward asymmetry at the ±1 lags was not different between the two states.

As for prior Studies, we conducted a Bayesian model comparison in which we compared the three-way ANOVA reported above to a null model without an attentional state variable. Once again, this analysis revealed very strong evidence in favor of the null model (BF > 150).

We therefore once again replicated the finding that recall is temporally organized. However, the temporal organization of recall was not different between the two attentional states. This replicates the null findings from the lag-CRP analyses in Studies 1–3. This null effect was observed even though overall recall was higher for items encoded "in the zone" vs "out of the zone".

5.3.5. *No differences in event transition types between the two attentional states*

We next examined recall transitions as a function of the type of event segment (see Fig. 2 and Section 2.1.2.5).

To do this, we performed a two-way repeated-measures ANOVA with transition type (3 levels) and attentional state ("in the zone" vs. "out of the zone") as factors. As before, we hypothesized that "same state, same event" and "same state, different event" transitions may be more likely for items encoded "in the zone" vs. "out of the zone".

We found a main effect of transition type ($F_{1.09,\ 134.05} = 30.45$, $p < 0.0001$, $\eta_p^2 = 0.20$, BF < 1/150). The main effect of attentional state ($F_{1,123} = 1.23$, $p = 0.27$, $\eta_p^2 = 0.01$, BF = 6.93), and the interaction between attentional state and transition type ($F_{1.14,\ 140.14} = 0.36$, $p = 0.58$, $\eta_p^2 = 0.003$, BF = 25.54), were not statistically significant. This suggests that each type of recall transition is not differentially likely for items encoded "in the zone" and "out of the zone" (Fig. 14).

We conducted follow-up *t*-tests to understand the main effect of transition types, collapsing across the two attentional states. As before, we found that "same state, same event" transitions were significantly more likely compared to "same state, different event" transitions ($t_{123} = 5.75$, $p < 0.0001$, Cohen's dz. = 0.52, 95% CI [0.002, 0.004]) and "different state, different event" transitions ($t_{123} = 5.43$, $p < 0.0001$, Cohen's dz. = 0.49, 95% CI [0.002, 0.004]), There was no significant difference between "different state, different event" transitions and "same state, different event" transitions ($t_{123} = 1.46$, $p = 0.15$, Cohen's dz. = 0.13, 95% CI [−0.00007, 0.0005]).

As for prior Studies we conducted a Bayesian model comparison in which we compared the two-way ANOVA reported above to a null model without an attentional state variable. This analysis once again revealed very strong evidence in favor of the null model (BF > 150). Thus, as in prior Studies, we found no evidence that recall transitions differed between the two attentional states.

5.3.6. *Secondary analysis: faster vs. slower RTs*

As before, we examined task performance and subsequent memory with a secondary analysis in which encoding RTs were divided, by a median split, into those that were faster vs. those that were slower. Participants made more errors during attentional states characterized by

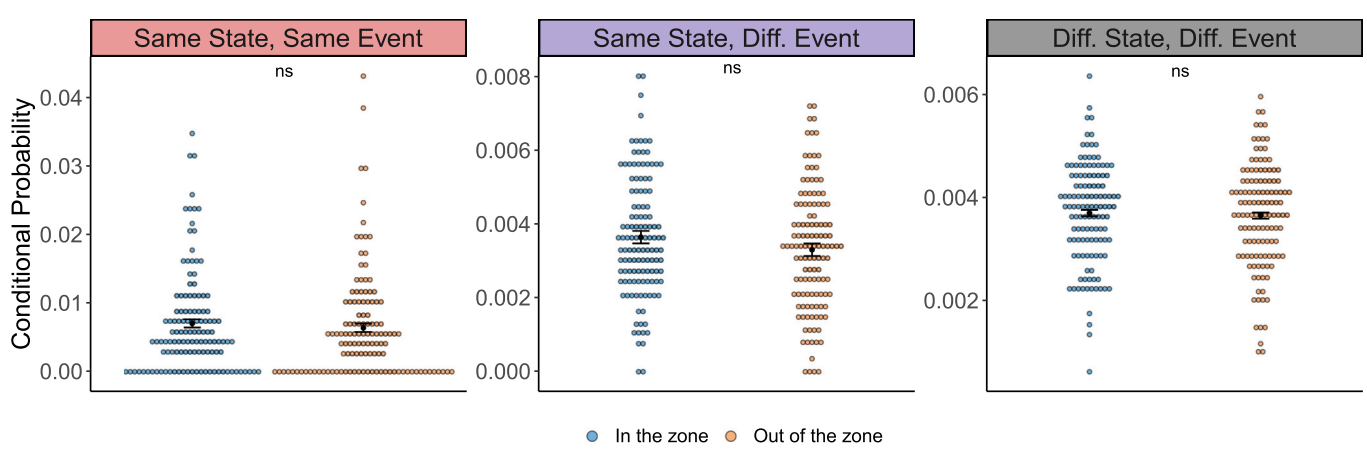

Fig. 14. Recall transitions as a function of event type in Study 4. Recall transitions are shown based on whether they occurred within an "event segment' of a particular attentional state (same state, same event), across event segments of a given attentional state (same state, different event), or between attentional states (different state, different event; see Fig. 2**)**. There were no significant differences in any transition type between the two attentional states. Individual points indicate the conditional probability of each transition type (i.e., the number of times each transition type occurred divided by the number of opportunities to make a transition of that type) for each individual, separately for items encoded "in the zone" and "out of the zone". Black dots indicate the mean; error bars indicate the standard error of the within-participant difference between "in the zone" and "out of the zone". ns = not statistically significant.

faster vs. slower RTs ($t_{123} = 2.04$, $p = 0.043$, Cohen's dz. = 0.18, 95% CI [0.02, 1.06], BF = 1.35). They also exhibited worse subsequent recall for items associated with faster vs. slower RTs at encoding ($t_{123} = 2.18$, $p = 0.03$, Cohen's dz. = 0.20, 95% CI [0.09, 1.87], BF = 1.02). These results are conceptually similar to those with the VTC analysis, in which more errors and worse recall were linked to "out of the zone" vs. "in the zone" states.

The lag-CRP analysis, however, failed to reveal any differences in the temporal organization of recall: there was only a main effect of absolute lag ($F_{15.56,\ 1914.03} = 6.46$, $p < 0.0001$, $\eta_p^2 = 0.049$, BF < 1/150; all other ps > 0.24; all other BFs > 26.92).

The event segment analysis revealed a main effect of transition type ($F_{1.03,126.94} = 32.68$, $p < 0.0001$, $\eta_p^2 = 0.21$, BF < 1/150), reflecting significantly more "same state, same event" transitions compared to "same state, different event" transitions ($t_{123} = 5.72$, $p < 0.0001$, Cohen's dz. = 0.51, 95% CI [0.003, 0.007]) and "different state, different event" transitions ($t_{123} = 5.78$, $p < 0.0001$, Cohen's dz. = 0.52, 95% CI [0.003, 0.007]). There was no significant difference between "different state, different event" and "same state, different event" transitions ($t_{123} = 0.74$, $p = 0.46$, Cohen's dz. = 0.07, 95% CI [−0.0002, 0.0004]). There was no main effect of attentional state ($F_{1,123} = 0.085$, $p = 0.77$, $\eta_p^2 = 0.0007$, BF = 11.26) nor an attentional state by transition type interaction ($F_{1.05,129.40} = 0.065$, $p = 0.81$, $\eta_p^2 = 0.0005$, BF = 32.97).

Finally, as for prior Studies, we performed a Bayesian model comparison between the lag-CRP model reported above and a null model without an attentional state variable. This comparison revealed very strong evidence in favor of the null model (BF > 150). Likewise, the same approach for the event segment analysis revealed very strong evidence in favor of the null model (BF > 150).

Thus, as for other Studies, there was no statistically significant difference in recall organization between attentional states associated with faster vs. slower RTs at encoding. Our main conclusions — lack of evidence that attentional states at encoding influence the temporal organization of memory — therefore hold with this alternative characterization of attentional states.

5.4. Discussion

In Study 4, we sought to replicate the findings from our prior Studies, particularly the lack of a difference in the temporal structure of recall for items encoded "in the zone" vs. "out of the zone". We used a design similar to Study 2 but increased block length and reduced stimulus presentation time to make our design more similar to the traditional gradCPT. As in our other Studies, we failed to find any evidence for more temporally structured recall for "in the zone" vs "out of the zone" encoding states.

Interestingly though, this was the first Study in which we found an effect of "in the zone" vs. "out of the zone" attentional states on subsequent recall: participants recalled more items encoded during an "in the zone" state compared to an "out of the zone" state. Thus, although we found fewer online errors and better overall recall for items encoded "in the zone" vs. "out of the zone" (Fig. 12), and although we replicated temporal contiguity effects in overall recall, we still failed to find evidence for differences across attentional states in recall organization. This suggests that our VTC analysis was able to successfully differentiate between better and worse attentional states, but these states were remarkably similar in the temporal organization of recall. We discuss the implications of our findings and their relation to prior work in the **General Discussion**.

6. General discussion

6.1. Summary of findings

We examined the behavioral effects of endogenous fluctuations in attention on the temporal organization of memory. We used response time variability at encoding to characterize two attentional states: the relatively good "in the zone" state and the relatively worse "out of the zone" state. We hypothesized that good (vs. bad) attentional states at encoding will be more conducive to maintaining temporal context representations, thus promoting more temporally organized recall and facilitating "leaps" between temporally distant but cognitively similar attentional states. However, across four Studies we failed to find evidence to support either hypothesis. Indeed, Bayesian comparisons — between models that included attentional state as a variable and null models that did not — consistently revealed strong evidence for the null hypothesis that attentional states at encoding do not influence the temporal organization of recall.

We replicated previous findings that individuals make more errors in online task performance during "out of the zone" states (e.g., Esterman et al., 2013, 2014; Fortenbaugh et al., 2018; Rosenberg et al., 2013; Rosenberg et al., 2015). In Study 4, we also found that recall was worse when encoding occurred in an "out of the zone" state. We also replicated several well-established memory phenomena, including temporal contiguity effects and forward asymmetry in recall (Healey et al., 2019; Howard & Kahana, 2002b; Kahana, 1996). Despite this, we found no evidence that the temporal organization of recall was affected by attentional fluctuations at encoding: recall was robustly temporally organized, even when encoding occurred in relatively poor attentional states. Furthermore, even when we conducted an analysis that combined Studies 1, 2, and 4 (our diagnostic experiments, in which "out of the zone" attentional states were associated with more online errors), we found no evidence of differential temporal organization of recall as a function of encoding attentional state (all *ps* > 0.29 for interactions involving attentional state). Yet, there was strong evidence for temporally organized recall generally (main effect of lag: $F_{11.22,2871.99} = 24.78$, $p < 0.0001$, $\eta_p^2 = 0.09$; lag by direction interaction: $F_{11.40,2918.07} = 2.49$, $p = 0.0036$, $\eta_p^2 = 0.01$). Together, our findings suggest that temporal context serves as a strong scaffold for episodic memory, one that can overcome spontaneous fluctuations in attentional states. Furthermore, our Studies highlight the difficulty of merging tasks of sustained attention and memory recall, and the numerous factors that must be considered when doing so (e.g., list length, trial-unique items, repetitive responses). We explore other potential reasons for our findings, their implications, and recommendations for future research in the sections below.

6.2. Exploring reasons for the null effect of attentional states on recall organization

Why did we not see the hypothesized relationship between attentional states and temporal organization of recall? One possibility is that converting our experiments to online studies increased the noise in our data, hence obscuring any potential effects. However, data from Study 1 suggests this isn't the case: in control analyses, we found no differences in any measures of interest between online and in-person participants. Furthermore, we replicated established in-lab phenomena in our online-only studies (Studies 2–4) such as the temporal contiguity effect and forward asymmetry in free recall, as well as more errors for "out of the zone" attentional states. Thus, it is unlikely that moving to online experiments was the main reason behind the lack of evidence supporting our hypothesis.

A second possibility is that measures of RT variability are not sensitive to spontaneous fluctuations in attentional states, and thus, we failed to characterize these states. However, there is strong evidence from sustained attention studies that response time variability effectively captures subtle fluctuations in attentional states, which can then be related to online task performance (e.g., Esterman et al., 2013, 2014; Robertson et al., 1997; Rosenberg et al., 2013). Furthermore, other studies have used RT variability-based attentional states as a trait-level measure and related it to episodic memory (Madore et al., 2020). Our

results were consistent with these effects: compared to "in the zone" states, "out of the zone" attentional states were associated with more errors during the encoding task and, in Study 4, worse recall. Furthermore, we found that RTs reliably fluctuated around errors, and that RT variability tended to increase over the course of the experiment. These findings together suggest that RTs, and RT variability, capture important cognitive phenomena in our tasks.

Despite the success of the variance time course (VTC) analysis that we focused on in the current study (Esterman et al., 2013, 2014; Madore et al., 2020; Rosenberg et al., 2013; Rosenberg et al., 2015), there are alternative ways to quantify better vs worse attentional states. For example, some studies of sustained attention have shown that faster (vs. slower) RTs are linked to more online errors and worse subsequent recognition memory (e.g., Cheyne et al., 2006; deBettencourt et al., 2018; McVay & Kane, 2012; Robertson et al., 1997; Wakeland-Hart et al., 2022). Conversely, other studies link longer eye fixations and slower response times to mind wandering and attentional lapses (e.g., Smallwood et al., 2003; Cheyne et al., 2009; Feng et al., 2013; Yanko & Spalek, 2013; Kam & Handy, 2014; Henríquez et al., 2016; Krasich et al., 2018; Zhang et al., 2021). Thus, some studies suggest that slower (vs. faster) RTs are associated with better attentional states; other studies suggest the opposite. Both of these approaches can be contrasted to the VTC analysis, in which RTs that are too fast or too slow (relative to the mean) are considered to reflect a poor attentional state. To test whether an alternative characterization of RT-based attentional states yields different results, we performed a secondary analysis for each Study. We compared online task performance and subsequent memory as a function of faster RTs vs. slower RTs (from a median split) during encoding.

We found that this approach was generally inferior to the VTC analysis, in that it could not reliably detect differences in online errors based on attentional state. The strongest support for this approach came in Study 4, for which there was a significant effect for online errors (more errors for attentional states associated with faster vs. slower RTs) and a significant effect on recall (worse recall for attentional states associated with faster vs. slower RTs). Across all Studies, however, this approach — like the VTC analysis — failed to reveal differences in the temporal organization of recall based on attentional states at encoding. Thus, our choice of the VTC analysis over this alternative approach does not change our main conclusions. Nevertheless, other measures of attentional fluctuations, such as pupil diameter changes linked to physiological arousal (Clewett, Gasser, & Lila, 2020; Decker et al., 2020; Unsworth et al., 2018; van den Brink, Murphy, & Nieuwenhuis, 2016; Zhang, Zhang, & Xu, 2020), could be used in future studies to link attentional fluctuations to the temporal organization of recall.

A third possibility is that, in our Studies, attentional fluctuations had a more minor effect on recall than other variables did. For example, recall can also be structured by the semantic similarity of encoded items (Healey et al., 2019; Howard & Kahana, 2002a; Polyn et al., 2009). Individuals likely used semantic information to guide recall in our Studies, but we nevertheless observed consistent and reliable temporal structure in recall as well: every Study showed a robust temporal contiguity effect in recall. This is in line with prior work demonstrating that, while many variables — such as list length, presentation times, incidental vs. intentional encoding, emotional salience, and inter-item distraction — can influence the the magnitude of the temporal contiguity effect, it tends to be reliably present (Dester, Lazarus, Uitvlugt, & Healey, 2020; Healey et al., 2019; Lazarus, Dester, Uitvlugt, & Healey, 2020; Peris-Yague, Frank, & Strange, 2021).

Future studies could determine if semantic clustering differs for items encoded "in the zone" vs. "out of the zone". If so, attentional modulation of semantic but not temporal clustering would provide additional evidence for dissociations between these aspects of recall organization (Healey et al., 2019; Howard & Kahana, 2002a; Polyn et al., 2009) and provide further constraints for models of memory.

Recall tests only allow assessment of memories that are sufficiently strong as to be brought to mind without external cues. Thus, one interesting potential mechanism for the general lack of difference in overall recall between attentional states is that the items that were recalled were particularly distinctive, memorable, or salient to a given participant. This may make memory for those items less susceptible to attentional fluctuations than they otherwise would be. We used lists of relatively well-known objects to try to minimize large differences in salience or distinctiveness, but this does not rule out that some objects may be more salient or distinctive for any given participant or more memorable generally (Bainbridge, 2019, 2020; Bainbridge, Isola, & Oliva, 2013; Isola, Xiao, Torralba, & Oliva, 2011). However, the presence of more vs. less memorable items does not necessarily explain similar levels of *temporal structure* in recall across attentional states. Because items were randomly ordered for every participant, particularly memorable items were unlikely to be temporally clustered in a similar way across "in the zone" and "out of the zone" states. That said, a compelling hypothesis is that items that are more distinctive or otherwise memorable "stick" in memory and can overcome periods of low attention to support not only their recall but also access to the surrounding temporal context, and thus recall of adjacent items. This mechanism could be explicitly tested in future work inspired by our findings. Prior relevant work has shown that item memorability and attentional states independently predict recognition memory (Wakeland-Hart et al., 2022), but such work could be extended to determine how attentional states and item memorability interact to support overall recall and the temporal structure of recall.

Moreover, other studies could use alternative measures of capturing the temporal structure of memory that do not depend on free recall (e.g., recognition or cued recall), and thus allow assessment of weaker memories. For instance, a study by Schwartz, Howard, Jing, and Kahana (2005), used a temporally structured recognition memory task to examine the temporal organization of memory. They showed that, when individuals recognized a scene with high confidence, the probability that the next scene would also be recognized with high confidence decreased as the encoding distance between those scenes increased. Thus, future work can index attentional fluctuations during the encoding phase of such a task, and relate these fluctuations to subsequent recognition memory and its sensitivity to the temporal structure of the test. This would allow examination of how attentional fluctuations influence memories that are too weak to be recalled but can nevertheless be recognized. Indeed, other studies have found that attentional fluctuations at encoding influence recognition memory overall (deBettencourt et al., 2018; Wakeland-Hart et al., 2022), leaving open the possibility that such fluctuations also influence the temporal structure of recognition memory.

Finally, it is worth noting that our Studies deviated from standard list-learning recall tasks (Healey & Kahana, 2014; Kahana, 1996; Sederberg et al., 2010) in several ways. For example, we used images instead of words; we had longer lists of items and fewer lists, and we used gradual transitions between images (in Studies 1, 2, and 4). The changes we made were, however, necessary to address our question. Images, and gradual transitions between them, were needed to align with the traditional gradCPT approach, which was used to validate the VTC analysis and identification of "in the zone" vs. "out of the zone" states (Esterman et al., 2013, 2014; Esterman et al., 2015; Esterman, Poole, Liu, & DeGutis, 2017; Esterman & Rothlein, 2019; Fortenbaugh et al., 2015, 2018; Rosenberg et al., 2013; Rothlein, DeGutis, & Esterman, 2018; Yamashita et al., 2021). Images are used in several studies of recall organization (e.g., Nguyen & McDaniel, 2015; Clark & Bruno, 2021; Kelly & Beran, 2021; also see Healey et al., 2019); other studies have also examined recall organization for naturalistic event memory (e.g., Diamond & Levine, 2020; Moreton & Ward, 2010; Uitvlugt & Healey, 2019). As in those studies, we replicated typical properties of lag-CRP curves in our experiments. Thus, our use of images is unlikely to have led to qualitative differences in how recall operated. We also removed the gradual transitions between images in Study 3, but still failed to find a link between attentional states and recall organization.

Our lists were longer than those typically used in list-learning studies because blocks had to be long enough to induce attentional fluctuations; attention is unlikely to robustly fluctuate in a shorter list of 10–25 items. Likewise, we had fewer lists because individuals encoded more items within each list. Despite these departures from list-learning studies, however, we were able to replicate typical properties of lag-CRP curves. Even though, in our experiments, absolute lag-CRP values at nearby lags were lower than those in other studies (e.g., Healey & Kahana, 2014; Kahana, 1996; Sederberg et al., 2010), the typical shape of the lag-CRP curve, and effects of temporal contiguity and forward asymmetry, were observed. Indeed, previous work has noted that, although lag-CRP values are lower for longer lists, both short and long lists show clear temporal contiguity effects (Healey et al., 2019). This highlights the reliability and robustness of temporal organization in recall (Healey et al., 2019). Critically, as noted above, these deviations from standard procedures were not random, but were necessary for us to test our hypotheses. Because we were able to replicate not only typical lag-CRP findings, but also findings from the sustained attention literature (Esterman et al., 2014; Esterman et al., 2015; Esterman et al., 2017; Esterman & Rothlein, 2019; Fortenbaugh et al., 2018; Kucyi, Esterman, Riley, & Valera, 2016; Kucyi, Hove, Esterman, Hutchison, & Valera, 2017; Rosenberg et al., 2015; Rothlein et al., 2018; Yamashita et al., 2021), we believe it is unlikely that our design decisions qualitatively altered attention and memory as studied in those procedures.

6.3. Suggestions for future research

An important contribution of this research is that it highlights the difficulty of balancing the demands of measuring sustained attention and recall. Taxing sustained attention requires a long, boring, and repetitive task — task features that work against robust recall for the encountered items. Thus, the ideal task for addressing how attentional fluctuations affect the temporal structure of recall must be long enough to induce lapses in attention but not so long that memory performance is at floor.

One key strength of the multiple experimental designs used across our four Studies is in building confidence in the null effect of attentional state on the temporal organization of recall: our lists lengths varied from 30 to 120 items, our tasks varied the encoding question, and we tested both gradual and abrupt image onsets. Thus, our findings should be useful for the field, particularly for researchers who are concerned about potential confounds of fluctuating attention during memory encoding and how that may affect recall organization. Future research on this topic would also benefit from exploring myriad experimental procedures to ensure that the results are not specific to any given paradigm.

Given our null effect of attentional state on recall organization across the many different experimental designs we used, we believe it is unlikely that attentional fluctuations in standard list-learning tasks reliably influence the temporal organization of recall within a list. This is because encoding lists in those tasks are likely too short to elicit robust attentional fluctuations. Nevertheless, there are multiple alternative ways that future research can assess the relationship between attentional fluctuations and memory organization, which we summarize here. We hope these strategies may help researchers address similar questions while avoiding the task complexities that we encountered.

First, our approach focused on assessing attentional fluctuations within long lists of items, in keeping with the standard approach in sustained attention research. But it is likely that attention also fluctuates slowly over the course of an entire experimental session (indeed, we found that RTs became progressively more variable across the session). Slow changes in attentional lapsing may influence recall organization from early to late in the experiment. Researchers could capitalize on such slower fluctuations by implementing the variance time course (VTC) analysis over RTs from the entire experiment rather than within a list, which has the benefit of allowing researchers to use many short lists (e.g., 15 lists of 16 items, or similar) as is often done in standard list-learning tasks (e.g., Murdock Jr., 1962; Kahana, 1996; Sederberg et al., 2010; Cortis et al., 2015; also see Healey et al., 2019). This approach would allow researchers to determine if most "in the zone" states are early on in the experimental session rather than later, and whether that is related to more temporal organization for lists encoded early vs. late in the experiment. Alternatively, researchers could use pupil measures of attentional states (Clewett et al., 2020; Decker et al., 2020; Kang, Huffer, & Wheatley, 2014; Unsworth et al., 2018; Unsworth & Robison, 2016; van den Brink et al., 2016; Zhang et al., 2020) to measure fluctuations in attention across an experimental session, and determine if that is related to reduced temporal clustering in later lists vs. earlier lists.

Researchers could also use tests of cued recall or temporally structured recognition tests, such as the recognition approach reviewed earlier (Schwartz et al., 2005). Such tests allow assessments of temporal structure in memory while also probing weaker memories that may not be freely recalled. For example, researchers could determine the likelihood that, given an encoded item as a cue, the participant recalls the next item encoded (lag +1) as opposed to the previous encoded item (lag −1) or items encoded further away (lags >1). These approaches overcome the limitation that only relatively strongly encoded items are freely recalled, and such items may be salient or interesting enough to "survive" periods of low attention. Such approaches could also be combined with measures of item distinctiveness or memorability (Wakeland-Hart et al., 2022; see Section 6.2: **Exploring reasons for the null effect of attentional states on recall organization**), allowing researchers to determine separate and interacting effects of attentional states and memorability. For example, item memorability may increase recall likelihood more for items encountered during poor vs. good attentional states. Furthermore, more memorable items may aid in the recovery of temporal context and trigger recall of the subsequently encoded item when they are presented as a cue in a cued recall task.

Finally, the goal of our project was to examine how attentional fluctuations influence the *temporal* organization of recall, leaving open the question of how such fluctuations influence the *semantic* clustering of recall. Thus, all of the above approaches could be repeated with semantic rather than temporal clustering as the critical output measure. Together, these approaches would yield insight into how various measures of attentional fluctuations (response time vs. pupil measures, within-list vs. across-experiment) influence different types of recall organization (temporal vs. semantic). Such work could lead to a mechanistic understanding of how various factors can overcome poor attentional states to support structured memories.

6.4. Relation to prior work

While we did not find evidence supporting our hypothesis that spontaneous fluctuations in attention influence the temporal organization of memory, it is likely that stronger manipulations of attention would affect recall organization. Indeed, there is ample work showing that attention influences memory (see Aly & Turk-Browne, 2017; Chun & Turk-Browne, 2007). Studies involving dual tasks show that divided attention at encoding is associated with worse memory at retrieval (for e.g. Baddeley, Lewis, Eldridge, & Thomson, 1984; Craik et al., 1996; Uncapher & Rugg, 2005). Other tasks involving experimenter-manipulated attention also show robust effects on memory (for example, Aly & Turk-Browne, 2016; Uncapher & Rugg, 2009; Yi & Chun, 2005). One recent study showed that the type of attention at encoding (whether there was a semantic task or no task) influences temporal clustering in recall (Long & Kahana, 2017). Another recent study showed that introducing distractions during encoding disrupts the temporal contiguity effect (Cutler, Jeon, & Polyn, 2020). Thus, experimentally manipulated attention robustly impacts memory performance generally as well as the temporal organization of recall.

There is evidence that spontaneous attentional fluctuations influence memory, but these studies do not examine the temporal dynamics of

recall. One such body of work is research on mind wandering. In these studies, participants are asked to report whether they are "on-task" or "off-task" at various intervals (for example, Smallwood et al., 2003, 2008; Metcalfe & Xu, 2016; Xu & Metcalfe, 2016; Xu et al., 2018; also see Smallwood & Schooler, 2015; & Christoff et al., 2016). These studies have found that more mind wandering is associated with decreased change detection in memory (Garlitch & Wahlheim, 2020), less precise cued recall (Martarelli & Ovalle-Fresa, 2021), and worse memory performance in general (Risko et al., 2012; Smallwood et al., 2003). Furthermore, as mentioned above, there is evidence that worse attentional states at encoding, as indexed by RTs, are associated with worse recognition memory (deBettencourt et al., 2018; Decker et al., 2020; Wakeland-Hart et al., 2022) and worse associative memory (Elshiekh & Rajah, 2021) in a subsequent test.

Together, the above-reviewed research shows that experimental manipulations of attention have effects on the temporal structure of recall and that spontaneous attentional fluctuations have effects on other forms of memory. Our current findings suggest that temporal context may be a more powerful driver of memory organization than attentional fluctuations — at least those that occur naturally and spontaneously during a task. Determining the conditions in which spontaneous attentional fluctuations may influence the temporal structure of memory requires further work; for example, future studies using temporally structured recognition tests or alternative measures of attentional fluctuations (e.g., pupil diameter) may yield important insights.

6.5. Conclusion

Across four Studies, we did not find any evidence that attentional fluctuations during encoding, as measured by the speed or variability of response times, influenced the temporal organization of recall. Temporal organization of recall is robust, even for memories encoded during relatively poor attentional states. These findings suggest that temporal context serves as a strong scaffold for episodic memory, one that can overcome spontaneous fluctuations in attentional states. Future research can assess the generality of these results by examining other measures of attention and memory.

CRediT authorship contribution statement

Manasi Jayakumar: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Chinmayi Balusu:** Data curation, Investigation, Writing – review & editing. **Mariam Aly:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

Deidentified data, materials, and code are available at https://osf.io/hs25k. A self-contained computing environment with data and code is also available on Code Ocean at https://doi.org/10.24433/CO.3162457.v1 for reproducing the main analyses, results, and figures.

Acknowledgments

We would like to thank Chris Baldassano, the Alyssano group, Lila Davachi, the Davachi Memory Lab, Paul Bloom, Monica Thieu, Alexandra Decker, Jeremy Manning, Paxton Fitzpatrick, Mike Esterman, and Megan DeBettencourt for valuable insights on this project and advice about analyses. We would also like to thank Matthew Siegelman for the OptSeg tool used in Study 1 and Somasundaram Ardhanareeswaran for help with automating recall transcriptions. This work was funded by an NSF CAREER Award (BCS-1844241) to M.A.

Appendix A. Supplementary table

Supplementary table to this article can be found online at https://doi.org/10.1016/j.cognition.2023.105408.

References

Aly, M., & Turk-Browne, N. B. (2016). Attention promotes episodic encoding by stabilizing hippocampal representations. *Proceedings of the National Academy of Sciences, 113*(4), E420–E429. https://doi.org/10.1073/pnas.1518931113

Aly, M., & Turk-Browne, N. B. (2017). How hippocampal memory shapes, and is shaped by, attention. In *The Hippocampus from cells to systems: Structure, connectivity, and functional contributions to memory and flexible cognition* (pp. 1–589). https://doi.org/10.1007/978-3-319-50406-3

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: I. evidence from divided attention costs. *Psychology and Aging, 13*(3), 405–423. https://doi.org/10.1037/0882-7974.13.3.405

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Baddeley, A., Lewis, V., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*(4), 518–540. https://doi.org/10.1037/0096-3445.113.4.518

Bainbridge, W. A. (2019). Chapter one - memorability: How what we see influences what we remember. In K. D. Federmeier, & D. M. Beck (Eds.), *70. Psychology of learning and motivation* (pp. 1–27). Academic Press. https://doi.org/10.1016/bs.plm.2019.02.001.

Bainbridge, W. A. (2020). The resiliency of image memorability: A predictor of memory separate from attention and priming. *Neuropsychologia, 141*, Article 107408. https://doi.org/10.1016/j.neuropsychologia.2020.107408

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General, 142*, 1323–1334. https://doi.org/10.1037/a0033872

Bainbridge, W. A., & Oliva, A. (2015a). Interaction envelope: Local spatial representations of objects at all scales in scene-selective regions. *Neuroimage, 122*, 408–416. https://doi.org/10.1016/j.neuroimage.2015.07.066

Bainbridge, W. A., & Oliva, A. (2015b). A toolbox and sample object perception data for equalization of natural images. *Data in Brief, 5*, 846–851. https://doi.org/10.1016/j.dib.2015.10.030

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2020). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods, 26*(3), 295. https://doi.org/10.1037/met0000337

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7*(6), 543–554. https://doi.org/10.1177/1745691612459060

Brand, A., Bradley, M. T., Best, L. A., & Stoica, G. (2008). Accuracy of effect size estimates from published psychological research. *Perceptual and Motor Skills, 106*(2), 645–649. https://doi.org/10.2466/pms.106.2.645-649

van den Brink, R. L., Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil diameter tracks lapses of attention. *PLoS One, 11*(10), Article e0165274. https://doi.org/10.1371/journal.pone.0165274

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS One, 9*(9), Article e106953. https://doi.org/10.1371/journal.pone.0106953

Chan, S. C. Y., Applegate, M. C., Morton, N. W., Polyn, S. M., & Norman, K. A. (2017). Lingering representations of stimuli influence recall organization. *Neuropsychologia, 97*, 72–82. https://doi.org/10.1016/j.neuropsychologia.2017.01.029

Cheyne, J. A., Carriere, J. S. A., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and Cognition, 15*(3), 578–592. https://doi.org/10.1016/j.concog.2005.11.009

Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition, 111*(1), 98–113. https://doi.org/10.1016/j.cognition.2008.12.009

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience, 17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2010). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*(1), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology, 17*(2), 177–184. https://doi.org/10.1016/j.conb.2007.03.005

Clark, D. P., & Bruno, D. (2021). Time is of the essence: Exploring temporal and spatial organisation in episodic memory. *Quarterly Journal of Experimental Psychology, 74*(8), 1406–1417. https://doi.org/10.1177/1747021821993823

Clewett, D., & Davachi, L. (2017). The ebb and flow of experience determines the temporal structure of memory. *Current Opinion in Behavioral Sciences, 17*, 186–193. https://doi.org/10.1016/j.cobeha.2017.08.013

Clewett, D., Gasser, C., & Lila, D. (2020). Pupil-linked arousal signals track the temporal organization of events in memory. *Nature Communications, 11*(1). https://doi.org/10.1038/s41467-020-17851-9

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 1179–1214. https://doi.org/10.1037/xlm0000086

Craik, F. I. M., Eftekhari, E., & Binns, M. A. (2018). Effects of divided attention at encoding and retrieval: Further data. *Memory & Cognition, 46*(8), 1263–1277. https://doi.org/10.3758/s13421-018-0835-3

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The Effects of Divided Attention on\nEncoding and Retrieval Processes in Human Memory. *Journal of Experimental Psychology: General, 125*(2), 159–180. https://doi.org/10.1037/0096-3445.125.2.159

Cutler, R. A., Jeon, J., & Polyn, S. M. (2020, August). *Characterizing the interaction of temporal and semantic information in categorized memory search [poster session]*. Philadelphia, PA: Context and Episodic Memory Symposium.

Damaso, K., Williams, P., & Heathcote, A. (2020). Evidence for different types of errors being associated with different types of post-error changes. *Psychonomic Bulletin & Review, 27*(3), 435–440. https://doi.org/10.3758/s13423-019-01675-w

deBettencourt, M. T., Cohen, J. D., Lee, R. F., Norman, K. A., & Turk-Browne, N. B. (2015). Closed-loop training of attention with real-time brain imaging. *Nature Neuroscience, 18*(3), 470–475. https://doi.org/10.1038/nn.3940

deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review, 25*(2), 605–611. https://doi.org/10.3758/s13423-017-1309-5

Decker, A., Finn, A., & Duncan, K. (2020). Errors lead to transient impairments in memory formation. *Cognition, 204*, Article 104338. https://doi.org/10.1016/j.cognition.2020.104338

Dester, A. M., Lazarus, L. T. T., Uitvlugt, M. G., & Healey, K. M. (2020, August). A test of retrieved context models. In *Dynamics of recall after incidental encoding [poster session]*. Philadelphia, PA: Context and Episodic Memory Symposium.

Diamond, N. B., & Levine, B. (2020). Linking detail to temporal structure in naturalistic-event recall. *Psychological Science, 31*(12), 1557–1572. https://doi.org/10.1177/0956797620958651

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology: General, 142*(4), 1277. https://doi.org/10.1037/a0034024

DuBrow, S., & Davachi, L. (2014). Temporal memory is shaped by encoding stability and intervening item reactivation. *Journal of Neuroscience, 34*(42), 13998–14005. https://doi.org/10.1523/JNEUROSCI.2535-14.2014

DuBrow, S., & Davachi, L. (2016). Temporal binding within and across events. *Neurobiology of Learning and Memory*. https://doi.org/10.1016/j.nlm.2016.07.011

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift? *Current Opinion in Behavioral Sciences, 17*, 141–146. https://doi.org/10.1016/j.cobeha.2017.08.003

Dutilh, G., van Ravenzwaaij, D., Nieuwenhuis, S., van der Maas, H. L. J., Forstmann, B. U., & Wagenmakers, E.-J. (2012). How to measure post-error slowing: A confound and a simple solution. *Journal of Mathematical Psychology, 56*(3), 208–216. https://doi.org/10.1016/j.jmp.2012.04.001

Elshiekh, A., & Rajah, M. N. (2021). Trial-by-trial fluctuations in post-stimulus attention during memory encoding predict subsequent associative context memory performance. *BioRxiv*. https://doi.org/10.1101/2021.02.16.431454, 2021.02.16.431454.

Esterman, M., Liu, G., Okabe, H., Reagan, A., Thai, M., & DeGutis, J. (2015). Frontal eye field involvement in sustaining visual attention: Evidence from transcranial magnetic stimulation. *NeuroImage, 111*, 542–548. https://doi.org/10.1016/j.neuroimage.2015.01.044

Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex, 23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261

Esterman, M., Poole, V., Liu, G., & DeGutis, J. (2017). Modulating reward induces differential neurocognitive approaches to sustained attention. *Cerebral Cortex, 27*(8), 4022–4032. https://doi.org/10.1093/cercor/bhw214

Esterman, M., Rosenberg, M. D., & Noonan, S. K. (2014). Intrinsic fluctuations in sustained attention and distractor processing. *Journal of Neuroscience, 34*(5), 1724–1730. https://doi.org/10.1523/JNEUROSCI.2658-13.2014

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology, 29*, 174–180. https://doi.org/10.1016/j.copsyc.2019.03.005

Ezzyat, Y., & Davachi, L. (2010). What constitutes an episode in episodic memory? *Psychological Science*. https://doi.org/10.1177/0956797610393742

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron, 81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review, 20*(3), 586–592. https://doi.org/10.3758/s13423-012-0367-y

Fortenbaugh, F. C., DeGutis, J., Germine, L., Wilmer, J. B., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained attention across the life span in a sample of 10,000: Dissociating ability and strategy. *Psychological Science, 26*(9), 1497–1510. https://doi.org/10.1177/0956797615594896

Fortenbaugh, F. C., Rothlein, D., McGlinchey, R., DeGutis, J., & Esterman, M. (2018). Tracking behavioral and neural fluctuations during sustained attention: A robust replication and extension. *NeuroImage, 171*, 148–164. https://doi.org/10.1016/j.neuroimage.2018.01.002

Frank, D., Gray, O., & Montaldi, D. (2020). SOLID-similar object and lure image database. *Behavior Research Methods, 52*(1), 151–161. https://doi.org/10.3758/s13428-019-01211-7

Garlitch, S. M., & Wahlheim, C. N. (2020). The role of attentional fluctuation during study in recollecting episodic changes at test. *Memory & Cognition, 48*(5), 800–814. https://doi.org/10.3758/s13421-020-01018-4

Gelman, A., & Carlin, J. (2017). Some natural solutions to the p-value communication problem—And why they Won't work. *Journal of the American Statistical Association, 112*(519), 899–901. https://doi.org/10.1080/01621459.2017.1311263

Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General, 143*, 575–596. https://doi.org/10.1037/a0033715

Healey, M. K., Long, N. M., & Kahana, M. J. (2019). Contiguity in episodic memory. *Psychonomic Bulletin & Review, 26*(3), 699–720. https://doi.org/10.3758/s13423-018-1537-3

Henríquez, R. A., Chica, A. B., Billeke, P., & Bartolomeo, P. (2016). Fluctuating minds: Spontaneous psychophysical variability during mind-wandering. *PLoS One, 11*(2), Article e0147174. https://doi.org/10.1371/journal.pone.0147174

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(7), 1075. https://doi.org/10.1037/xlm0000503

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25*(4), 923–941. https://doi.org/10.1037/0278-7393.25.4.923

Howard, M. W., & Kahana, M. J. (2002a). When does semantic similarity help episodic retrieval? *Journal of Memory and Language, 46*(1), 85–98. https://doi.org/10.1006/jmla.2001.2798

Howard, M. W., & Kahana, M. J. (2002b). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Isola, P., Xiao, J., Torralba, A., & Oliva, A. (2011). What makes an image memorable? *Journal of Vision, 11*(11), 1282. https://doi.org/10.1167/11.11.1282

Jeffreys, H. (1961). *The theory of probability*. OUP Oxford.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition, 24*(1), 103–109. https://doi.org/10.3758/BF03197276

Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 530–540. https://doi.org/10.1037/0278-7393.28.3.530

Kam, J. W. Y., & Handy, T. C. (2014). Differential recruitment of executive resources during mind wandering. *Consciousness and Cognition, 26*, 51–63. https://doi.org/10.1016/j.concog.2014.03.002

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science, 21*(5), 348–354. https://doi.org/10.1177/0963721412454875

Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil dilation dynamics track attention to high-level information. *PLoS One, 9*(8), Article e102463. https://doi.org/10.1371/journal.pone.0102463

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kelly, A. J., & Beran, M. J. (2021). Multi-trial free recall dynamics in preschool children and young adults. *Journal of Cognitive Psychology, 33*(8), 837–852. https://doi.org/10.1080/20445911.2021.1977815

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*(6006), 932. https://doi.org/10.1126/science.1192439

Krasich, K., McManus, R., Hutt, S., Faber, M., D'Mello, S. K., & Brockmole, J. R. (2018). Gaze-based signatures of mind wandering during real-world scene processing. *Journal of Experimental Psychology: General, 147*(8), 1111. https://doi.org/10.1037/xge0000411

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science, 6*(3), 299–312. https://doi.org/10.1177/1745691611406925

Kucyi, A., Esterman, M., Riley, C. S., & Valera, E. M. (2016). Spontaneous default network activity reflects behavioral variability independent of mind-wandering. *Proceedings of the National Academy of Sciences, 113*(48), 13899–13904. https://doi.org/10.1073/pnas.1611743113

Kucyi, A., Hove, M. J., Esterman, M., Hutchison, R. M., & Valera, E. M. (2017). Dynamic brain network correlates of spontaneous fluctuations in attention. *Cerebral Cortex, 27*(3), 1831–1840. https://doi.org/10.1093/cercor/bhw029

LaRocque, J. J., Eichenbaum, A. S., Starrett, M. J., Rose, N. S., Emrich, S. M., & Postle, B. R. (2015). The short- and long-term fates of memory items retained outside the focus of attention. *Memory & Cognition, 43*(3), 453–468. https://doi.org/10.3758/s13421-014-0486-y

Lazarus, L. T. T., Dester, A. M., Uitvlugt, M. G., & Healey, K. M. (2020, August). *Modeling of reduced temporal contiguity for distinct items [poster session]*. Philadelphia, PA: Context and Episodic Memory Symposium.

Liu, C. C., & Aitkin, M. (2008). Bayes factors: Prior sensitivity and model generalizability. *Journal of Mathematical Psychology, 52*(6), 362–375. https://doi.org/10.1016/j.jmp.2008.03.002

Long, N. M., & Kahana, M. J. (2017). Modulation of task demands suggests that semantic processing interferes with the formation of episodic associations. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 43*(2), 167–176. https://doi.org/10.1037/xlm0000300

Madore, K. P., Khazenzon, A. M., Backes, C. W., Jiang, J., Uncapher, M. R., Norcia, A. M., & Wagner, A. D. (2020). Memory failure predicted by attention lapsing and media multitasking. *Nature, 587*(7832). https://doi.org/10.1038/s41586-020-2870-z. Article 7832.

Manning, J. R., & Kahana, M. J. (2012). Interpreting semantic clustering effects in free recall. *Memory, 20*(5), 511–517. https://doi.org/10.1080/09658211.2012.683010

Martarelli, C., & Ovalle-Fresa, R. (2021). In sight, out of mind? Effects of task disengagement on visual recall precision. *PsyArXiv*. https://doi.org/10.31234/osf.io/fdvzs

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "d'oh!": Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 525. https://doi.org/10.1037/a0025896

Metcalfe, J., & Xu, J. (2016). People mind wander more during massed than spaced inductive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(6), 978–984. https://doi.org/10.1037/xlm0000216

Moreton, B. J., & Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin & Review, 17*(4), 510–515. https://doi.org/10.3758/PBR.17.4.510

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods, 16*, 406–419. https://doi.org/10.1037/a0024377

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64*, 482–488. https://doi.org/10.1037/h0045106

Naveh-Benjamin, M., Craik, F. I., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1091.

Nguyen, K., & McDaniel, M. A. (2015). The picture complexity effect: Another list composition paradox. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 1026. https://doi.org/10.1037/xlm0000071

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition, 111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Palombo, D. J., Di Lascio, J. M., Howard, M. W., & Verfaellie, M. (2019). Medial temporal lobe amnesia is associated with a deficit in recovering temporal context. *Journal of Cognitive Neuroscience, 31*(2), 236–248. https://doi.org/10.1162/jocn_a_01344

Peris-Yague, A., Frank, D., & Strange, B. (2021, March). *Emotional and perceptual salience modulate conditional response probability curves in free recall [poster session]*. Cognitive Neuroscience Society Virtual Meeting.

Polyn, S. M., & Kahana, M. J. (2008). Memory search and the neural representation of context. *Trends in Cognitive Sciences, 12*(1), 24–30. https://doi.org/10.1016/j.tics.2007.10.010

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). Task context and organization in free recall. *Neuropsychologia, 47*(11), 2158–2163. https://doi.org/10.1016/j.neuropsychologia.2009.02.013

Purcell, B. A., & Kiani, R. (2016). Neural mechanisms of post-error adjustments of decision policy in parietal cortex. *Neuron, 89*(3), 658–671. https://doi.org/10.1016/j.neuron.2015.12.027

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology, 26*(2), 234–242. https://doi.org/10.1002/acp.1814

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). `oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747–758. https://doi.org/10.1016/S0028-3932(97)00015-8

Rosenberg, M. D., Finn, E. S., Constable, R. T., & Chun, M. M. (2015). Predicting moment-to-moment attentional state. *NeuroImage, 114*, 249–256. https://doi.org/10.1016/j.neuroimage.2015.03.032

Rosenberg, M. D., Finn, E. S., Scheinost, D., Constable, R. T., & Chun, M. M. (2017). Characterizing attention with predictive network models. *Trends in Cognitive Sciences, 21*(4), 290–302. https://doi.org/10.1016/j.tics.2017.01.011

Rothlein, D., DeGutis, J., & Esterman, M. (2018). Attentional fluctuations influence the neural Fidelity and connectivity of stimulus representations. *Journal of Cognitive Neuroscience, 30*(9), 1209–1228. https://doi.org/10.1162/jocn_a_01306

Rosenberg, M.D., Noonan, S., DeGutis, J., & Esterman, M. (2013). Sustaining visual attention in the face of distraction: A novel gradual onset continuous performance task. Attention Perception Psychopysics **75**, 426-439 (2013). https://doi.org/10.3758/s13414-012-0413-x.

Rouder, J. N., & Morey, R. D. (2011). A Bayes factor meta-analysis of Bem's ESP claim. *Psychonomic Bulletin & Review, 18*(4), 682–689. https://doi.org/10.3758/s13423-011-0088-7

Rouder, J. N., Morey, R. D., & Province, J. M. (2013). A Bayes factor meta-analysis of recent extrasensory perception experiments: Comment on storm, Tressoldi, and Di Risio (2010). *Psychological Bulletin, 139*, 241–247. https://doi.org/10.1037/a0029008

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Research Reviews, 35*(2), 146–160. https://doi.org/10.1016/S0165-0173(01)00044-3

Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science, 16*(11), 898–904. https://doi.org/10.1111/j.1467-9280.2005.01634.x

Sederberg, P. B., Miller, J. F., Howard, M. W., & Kahana, M. J. (2010). The temporal contiguity effect predicts episodic memory performance. *Memory & Cognition, 38*(6), 689–699. https://doi.org/10.3758/MC.38.6.689

Siegelman, M. (2019). Simulate an experiment with optimal stimuli segmentation. *GitHub*. https://github.com/msieg/OptSeg_Reproducible.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2016). *False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant* (p. 555). American Psychological Association. https://doi.org/10.1037/14805-033

Simonsohn, U. (2014). *Posterior-hacking: Selective reporting invalidates Bayesian results also (SSRN scholarly paper no. 2374040)*. https://doi.org/10.2139/ssrn.2374040

Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*(3), 452–484. https://doi.org/10.1016/S1053-8100(03)00018-7

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the Brain: Mind Wandering Reduces Cortical Analysis of External Events. *20*(3), 12.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin; US: American Psychological Association.*. https://doi.org/10.1037/0033-2909.132.6.946

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*(1), 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

Spillers, G. J., & Unsworth, N. (2011). Variation in working memory capacity and temporal–contextual retrieval from episodic memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1532–1539. https://doi.org/10.1037/a0024852

Troyer, A. K., & Craik, F. I. M. (2000). The effect of divided attention on memory for items and their context. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 54*(3), 161–171. https://doi.org/10.1037/h0087338

Troyer, A. K., Winocur, G., Craik, F. I. M., & Moscovitch, M. (1999). Source memory and divided attention: Reciprocal costs to primary and secondary tasks. *Neuropsychology, 13*(4), 467–474. https://doi.org/10.1037/0894-4105.13.4.467

Turk-Browne, N. B., Golomb, J. D., & Chun, M. M. (2013). Complementary attentional components of successful memory encoding. *NeuroImage, 66*, 553–562. https://doi.org/10.1016/j.neuroimage.2012.10.053

Uitvlugt, M. G., & Healey, M. K. (2019). Temporal proximity links unrelated news events in memory. *Psychological Science, 30*(1), 92–104. https://doi.org/10.1177/0956797618808474

Uncapher, M. R., Hutchinson, J. B., & Wagner, A. D. (2011). Dissociable effects of top-down and bottom-up attention during episodic encoding. *Journal of Neuroscience, 31*(35), 12613–12628. https://doi.org/10.1523/JNEUROSCI.0152-11.2011

Uncapher, M. R., & Rugg, M. D. (2005). Effects of divided attention on fMRI correlates of memory encoding. *Journal of Cognitive Neuroscience, 17*(12), 1923–1935. https://doi.org/10.1162/089892905775008616

Uncapher, M. R., & Rugg, M. D. (2009). Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *Journal of Neuroscience, 29*(25), 8270–8279.

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience, 16*(4), 601–615. https://doi.org/10.3758/s13415-016-0417-4

Unsworth, N., Robison, M. K., & Miller, A. L. (2018). Pupillary correlates of fluctuations in sustained attention. *Journal of Cognitive Neuroscience, 30*(9), 1241–1253. https://doi.org/10.1162/jocn_a_01251

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology, 100*, 426–432. https://doi.org/10.1037/a0022790

Wakeland-Hart, C. D., Cao, S. A., deBettencourt, M. T., Bainbridge, W. A., & Rosenberg, M. D. (2022). Predicting visual memory across images and within individuals. *Cognition, 227*, Article 105201. https://doi.org/10.1016/j.cognition.2022.105201

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience, 9*(7). https://doi.org/10.1038/nn1727. Article 7.

Williams, P., Heathcote, A., Nesbitt, K., & Eidels, A. (2016). Post-error recklessness and the hot hand. *Judgment and Decision making, 11*(2), 11.

Xu, J., Friedman, D., & Metcalfe, J. (2018). Attenuation of deep semantic processing during mind wandering: An ERP study. *Neuroreport, 29*(5), 380–384. https://doi.org/10.1097/WNR.0000000000000978

Xu, J., & Metcalfe, J. (2016). Studying in the region of proximal learning reduces mind wandering. *Memory & Cognition, 44*(5), 681–695. https://doi.org/10.3758/s13421-016-0589-8

Yamashita, A., Rothlein, D., Kucyi, A., Valera, E. M., Germine, L., Wilmer, J., … Esterman, M. (2021). Variable rather than extreme slow reaction times distinguish brain states during sustained attention. *Scientific Reports, 11*(1). https://doi.org/10.1038/s41598-021-94161-0. Article 1.

Yanko, M. R., & Spalek, T. M. (2013). Route familiarity breeds inattention: A driving simulator study. *Accident Analysis & Prevention, 57*, 80–86. https://doi.org/10.1016/j.aap.2013.04.003

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2011). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus, 21*(9), 968–979. https://doi.org/10.1002/hipo.20808

Yi, D.-J., & Chun, M. M. (2005). Attentional modulation of learning-related repetition attenuation effects in human Parahippocampal cortex. *Journal of Neuroscience, 25*(14), 3593–3600. https://doi.org/10.1523/JNEUROSCI.4677-04.2005

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*(2), 273. https://doi.org/10.1037/0033-2909.133.2.273

Zhang, H., Anderson, N. C., & Miller, K. F. (2021). Refixation patterns of mind-wandering during real-world scene perception. *Journal of Experimental Psychology: Human Perception and Performance, 47*(1), 36. https://doi.org/10.1037/xhp0000877

Zhang, W., Zhang, Y., & Xu, J. (2020). Performance and eye metrics correlates to out-of-the-zone sustained attention in GradCPT. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 64*(1), 112–116. https://doi.org/10.1177/1071181320641029