Psychonomic Bulletin & Review (2023) 30:1028–1040
https://doi.org/10.3758/s13423-022-02222-w

BRIEF REPORT

The spotlight turned inward: the time-course of focusing attention on memory

Gordon D. Logan[1] · Simon D. Lilburn[1] · Jana E. Ulrich[1]

Accepted: 16 November 2022 / Published online: 5 December 2022
© The Psychonomic Society, Inc. 2022

Abstract
Guided by the idea that memory retrieval is selective attention turned inward, we report four experiments examining the time-course of focusing attention on memory. We used a novel episodic flanker task that turns the famous perceptual flanker task inward, presenting memory lists followed by probes that asked whether a cued letter had appeared in the same position in the memory list. Like the perceptual flanker task, we manipulated distance to measure the sharpness of the focus of attention on memory, and compatibility to measure the resistance to distraction. To measure the time-course of focusing, we presented a cue indicating the probed position in the interval between the list and the probe and varied the interval between the cue and the probe (0, 250, 500, 750 ms). The main questions were whether the focus would become sharper and resistance to distraction would become stronger as cue–probe delay increased. Experiments 1a and 1b showed strong distance effects and strong cue–probe interval effects but no reliable interaction between them. Experiments 2a and 2b showed robust compatibility effects and cue–probe interval effects but no interaction between them. Thus, there is no evidence that the sharpness of the focus increases and little evidence that the resistance to distraction improves over time. The robust reduction in response time and slight increase in accuracy with cue–probe interval may reflect the time-course of orienting to the cued position in the memory list prior to focusing on the item it contains.

Keywords Attention · Memory · Time-course · Flanker · Compatibility

Introduction

The idea that memory retrieval is attention turned inward has driven research since James (1890) and is currently the topic of much investigation (Chun et al. 2011; Gazzaley & Nobre 2012; Kiyonaga & Egner 2013; Logan 2002). While much of the research has focused on capacity limitations, we focus on the selective nature of attention, investigating the time-course of focusing attention on a single cued item in memory (Souza & Oberauer, 2016) by varying the interval between the cue and the memory probe. The longer the interval, the more likely attention will be focused on the target item in memory, so the shorter the response time (RT; Eriksen & Hoffman, 1973; Logan, 2005; Souza & Oberauer, 2016; Sperling & Weichselgartner, 1995). We ask (1) Does the speedup in RT result from an increase in the sharpness of the focus of attention? (2) Does the speedup in RT reflect an increased ability to resist distraction?

We address these questions with an *episodic flanker task* (Logan et al., 2021) based on the Eriksen and Hoffman (1973) and Eriksen and Eriksen (1974) perceptual flanker task. The perceptual flanker task measures sharpness of the focus of attention by manipulating the distance between a cued target and flanking distractors (HSH vs. H S H). The target can appear in a randomly cued position (Eriksen & Hoffman, 1973), as in our task, or in a fixed, central position (Eriksen & Eriksen, 1974). RT is longer and accuracy is lower the closer the flankers. The flanker task measures resistance to distraction by manipulating compatibility between targets and distractors. Compatible targets and distractors point to the same response (HHH and SSS) and produce shorter RT and higher accuracy. Incompatible targets and distractors point to opposite responses (HSH and SHS) and produce longer RT and lower accuracy. Eriksen and colleagues explained distance and compatibility effects with a spotlight model, in which attention samples from a limited region of space centered on the target, and everything that

✉ Gordon D. Logan
gordon.logan@vanderbilt.edu

[1] Vanderbilt University, Nashville, TN, USA

falls within that region is sampled and competes to determine the response.

The episodic flanker task turns the spotlight of attention inward on memory and measures its properties using logic and models from the perceptual flanker task (Logan et al., 2021). Participants are given a list of six random letters to remember presented in a line (e.g., ABCDEF), followed by a probe display in which one of the letters is cued with a caret below it (^ represented here as an underline: ABCDEF). The task is to say whether the cued item in the probe occupied the same position in the memory list. The list structure provides a *distance metric* defined by selecting lures from different positions relative to the target (e.g., ##D### vs. ##E###). The relationship between nontarget items in the probe and the memory list manipulates *compatibility*, presenting probes with nontargets that are the *same* as (ABCDEF) or *different* from the memory list (e.g., GHCJKL). Same probes are compatible with "yes" responses (ABCDEF) and incompatible with "no" responses (ABEDCF). Different probes are incompatible with "yes" responses (GHCJKL) and compatible with "no" responses (GHEJKL). The compatibility effect is *episodic* because it depends on the relationship between the current memory list and the current probe. Different letters are compatible and incompatible on different trials.

Logan et al. (2021) replicated the classic distance and compatibility effects from the perceptual flanker task in the episodic flanker task, suggesting that the same spotlight of attention is used to retrieve items from perception and memory. The novel contribution of the present experiments is to study the time-course of focusing attention on memory, asking whether the focus becomes sharper and distraction is less effective as cue delay increases.

Logan et al. (2021)modeled the episodic flanker task by applying three established computational models of serial recall and interpreting their retrieval cues as spotlights of attention focused on memory. The *overlap* model (OVL) assumes noisy coding. Items are represented as distributions in memory space, and retrieval samples information from a region of space centered on one of the distributions (Estes, 1997; Logan, 1996) like a spotlight of attention (Eriksen & Hoffman, 1973; Posner, 1980). Distributions overlap, so samples include information from all the distributions that intrude in the sampled region. Compatible flankers support decisions about the target while incompatible flankers impair them. Nearby distributions contribute more than remote ones, explaining the distance effect. The *start–end* model (SEM) assumes items are associated with position codes and retrieved by probing memory with a position code, like object-based attention (Duncan, 1984; Kahneman et al., 1992). Position codes are more similar for adjacent items than for remote items, so adjacent items tend to be retrieved along with the target, explaining compatibility and distance effects. The *context retrieval and updating* model (CRU) assumes item coding. Items are associated with the current context at study and retrieved by probing stored contexts from the study list with the current context at test (Logan, 2021), like template-based attention (Bundesen, 1990; Logan, 2002). Nearby contexts are more similar than remote ones, explaining distance and compatibility effects. The retrieval decision is made by a racing diffusion process that predicts RT and accuracy. Each model has parameters that determine the sharpness of the focus and resistance to distraction (the standard deviation of the distributions in OVL, the steepness of the similarity gradient in SEM, and the updating parameter in CRU). Logan et al. found all three models fit distance and compatibility effects quantitatively. Here, we ask whether distance and compatibility effects change as attention focuses on memory. We report two sets of experiments, one addressing distance effects and one addressing compatibility effects.

Experiment 1: distance effects

The first experiment manipulated the distance between the cued position and the position that the cued item occupied in the memory list, ranging from 0 (item in original position, requiring a "yes" response) to 5 (original position 1–5 items away from cued position, requiring a "no" response). As in the flanker task, we assume the spotlight of attention is focused on the cued position in the memory list, and items near the cued position fall within the spotlight and influence processing. Lures that fall within the spotlight should activate a "yes" response, which interferes with the required "no" response, increasing RT and decreasing accuracy. The interference effect should decrease with distance because remote items are less likely to fall within the spotlight (Eriksen & Eriksen, 1974). In our models, sharpening the focus decreases activation of remote items and steepens the gradient of RT and accuracy around the cued position (Logan et al., 2021). The steepness of the gradient reflects the sharpness of the focus.

We ask whether the gradient becomes steeper as cue–probe interval increases and allows more time for attention to focus on the cued position in memory. Following zoom-lens (Eriksen & St. James, 1986), shrinking-spotlight (White et al., 2011), and dual-process (Cohen et al., 1992; Hübner et al., 2010) models of the flanker task, selectively attending to an item in memory might involve sharpening the focus over time. Alternatively, selectively attending might involve a preliminary process of orienting to the target in memory before focal processing begins (Smith & Ratcliff, 2009). Increasing cue–probe interval may provide more time for orienting but the focus may be the same for all intervals. The distance effect may not change over time.

Method

We ran two versions of Experiment 1 (and 2) that manipulated the same cue–probe intervals (0, 250, 500, 750 ms) in two ways. Experiment 1a fixed the interval between the list and the probe at 1,750 ms and presented cues 1,000–1,750 ms after the list. Experiment 1b fixed the interval between the list and the cue at 1,000 ms and presented probes 0–750 ms after the cue (1,000–1,750 ms after the list). We had no reason to expect differences between these procedures. We tested them both for generality and replication.

These experiments are replications of preliminary experiments we did that used more cue–probe intervals (0–700 ms in 100-ms steps) and fewer trials (480; one replication of the basic design for each cue–probe interval). There were 32 subjects in the original Experiment 1a and 31 in the original Experiment 1b. The results of the original experiments are reported in the Supplementary Information. The inferential statistics were clear and led to the same conclusions as the inferential statistics in the present experiments, but the data were noisy and the patterns were not as clear as we thought they should be, so we replicated the experiments with fewer cue–probe intervals (4) and more trials (720; three replications of the basic design for each cue–probe interval) to reduce experimental noise and produce more stable patterns.

Subjects We planned to test 32 subjects in each experiment. Logan et al. (2021) found distance effects in 32 of 32 subjects with the same manipulation, so we were confident we could replicate the distance effect. We also planned Experiments 1a and 1b as replications with a minor change in procedure, and that increased our confidence. Experiments 1a and 1b each included 32 subjects recruited online through Prolific (https://www.prolific.co/). Experiment 1a tested 32 subjects, and no subjects were excluded for failing to meet the accuracy criterion (described below). Experiment 1b tested 33 subjects and excluded one for failing to meet the accuracy criterion. Subjects who participated in one experiment were excluded from the other (and from Experiments 2a and 2b). Subjects matched on reported age (1a: $M = 30.2$ years, $SD = 5.8$ years; 1b: $M = 31.2$ years, $SD = 5.3$ years, one withheld) and had a similar gender distribution (1a: 17 males, 15 females; 1b: 29 males, six females, and one withheld). The eligibility criteria were set in Prolific to only include subjects between 18 and 40 years of age, located within the United States of America, who are native or fluent English speakers, with a rating of over 95%. Subjects completed the consent process in REDCap (https://www.project-redcap.org/). Each session was completed within 1.5 hours, and subjects were paid US$12 per hour. The study was approved by the Vanderbilt University Institutional Review Board.

Apparatus and stimuli The experiments were conducted online and run on subjects' personal computers. Subjects were instructed to use either Google Chrome or Mozilla Firefox to complete the experiment. Phone and tablet users were excluded from the Prolific intake, and the inputs required to progress within the experiment did not appear when using browsers in phones or tablets. The design of each session was generated and sent to each subject's computer using a custom Python backend. The experiment was controlled by JavaScript running within the web browser by using a custom function written to operate within jsPsych (de Leeuw, 2015). The memory lists consisted of six uppercase letters selected at random from the set of consonants (excluding vowels and *Y*). The probes consisted of five hash marks (#) and one probe letter in uppercase. The cue was an upward pointing caret (^) presented under the probed letter. Each position was cued equally often. The probed letter matched the letter in the same position on the memory list on half of the trials (distance = 0) and mismatched on the other half. Mismatching probes were selected from the five remaining positions in the list (distance = 1–5). All characters were presented in a monospaced typeface (Courier New or Courier if those fonts were installed on subjects' computers), displayed in white at 45 px in height. The background of the display was set to mid-gray ([127, 127, 127] in 24-bit RGB values).

Procedure The basic design required 60 trials to include all positions and distances. Probes appeared equally often in each probe position, and there were 10 trials for each probe position, five in which the probe contained the target letter and five sampling lures from the five remaining positions. With four cue–probe intervals, the complete design required 240 trials. We ran three replications of the complete design. The order of trials was randomized separately for each subject within each replication. The 720 trials were split into eight blocks (seven blocks of 96 trials with a single final block of 48 trials).

At the start of the experiment, the subject's web browser was instructed to enter into full-screen mode to reduce distractions from other applications. Both experiments consisted of 720 trials split into eight blocks (seven blocks of 96 trials with a single final block of 48 trials). The events on each trial are depicted in Fig. 1. Each trial began with a fixation cross presented in the center of the screen for 1,000 ms, followed by a six-letter memory list for 1,000 ms. Then, the screen was blanked for some time before a cue (^) appeared under the location of the letter to be judged and remained on throughout the cue–probe interval and the probe display, which remained on the screen until the subject responded. In Experiment 1a, the retention interval for all conditions was fixed at 1,750 ms, with the cue presented 0, 250, 500, or 750 ms prior to the probe display. In Experiment 1b, the

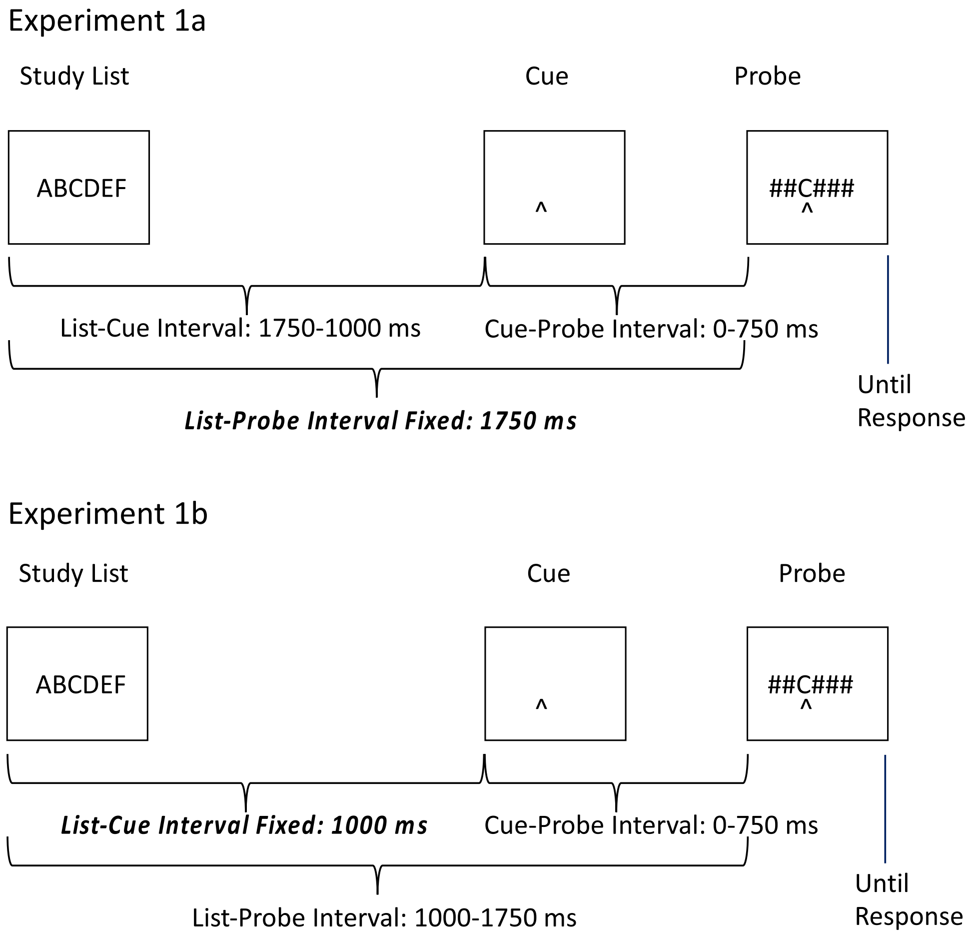

Fig. 1 Events on a trial in Experiments 1a and 1b. Each trial begins with a study list, followed after a list-cue interval by the cue, which remains on the screen until the response. The probe display appears after a cue–probe interval and remains on the screen until the response. Timing differs between experiments, as illustrated

interval between the memory display and the cue was fixed at a period of 1,000 ms, and the interval between the onset of the cue and the probe display (cue–probe interval) was 0, 250, 500, or 750 ms. Thus, the retention interval ranged from 1,000 to 1,750 ms. Each subject received the trials in a separate random order.

Subjects were instructed to indicate whether the cued letter in the probe was presented in the same position in the memory list, pressing the M key on the keyboard to indicate a "yes" response and the Z key to indicate a "no" response. The instructions were written and presented using a self-paced series of manually controlled slides. Subjects were allowed to review the instructions if they wished.

Subjects had to respond within 3,000 ms of the presentation of the probe list. If they took longer, the trial was terminated with the message "TOO SLOW" presented centrally in a red font for 3,000 ms. These trials were excluded from the analysis and treated as errors in calculating feedback during the task. At the end of each block, a screen was presented indicating the overall accuracy for the preceding block, and subjects were allowed to take a self-timed break. Every 5 minutes, the experiment checked whether accuracy was greater than 60%. If subjects fell below this accuracy criterion, they were warned to improve performance and given an opportunity to review the instructions. On the third warning, subjects were excluded from the experiment and from all subsequent analysis.

Data analysis In each experiment, we analyzed the RT and accuracy data with three planned contrasts. The first tested for the distance effect with linear contrast weights {2 1 0 −1 −2} for distances 1–5. Logan et al. (2021) showed that the distance effect was symmetrical for lures that preceded (negative distance) and followed (positive distance), so we calculated absolute distance rather than signed distance. The second contrast tested the cue–probe interval effect with linear contrast weights {3 1 −1 −3} for cue–probe intervals 0–750. The third contrast tested the predicted increase in the steepness (sharpness) of the distance effect with cue–probe interval, calculating the distance contrast for each cue–probe interval and then testing for an increase in the magnitude of the distance contrast with a linear contrast with weights $\{-3\ -1\ 1\ 3\}$ for cue–probe intervals 0–750. We used linear contrasts for each effect to capture the expected monotonic decrease in distance and cue–probe interval effects and the possible monotonic decrease in the distance effect with cue–probe interval. The linear contrasts do not capture the shape of the functions perfectly, but they capture the important downward trends with a single degree of freedom, and so provide concise tests of our hypotheses.

For each contrast, we divided the data for each subject into the relevant cells (1: five distances, 2: four cue–probe intervals, 3: 20 Distance × Cue–Probe intervals) and calculated the proportion of correct responses and the mean RT for correct responses with RT < 3,000 ms. Then we calculated the contrast for each individual subject, multiplying the mean RTs or the proportion of correct responses by the contrast weights and summing them. Then, we did a *t* test asking whether the mean contrast was significantly greater than zero. The error term was the standard error of the mean contrast value. We also counted the number of subjects who showed an effect in the expected direction and reported JZS Bayes Factors (BF) to quantify support for null (BF_{01}) and alternative (BF_{10}) hypotheses. To compare experiments, we performed independent samples *t* tests on the mean contrasts for RT and accuracy.

Results and discussion

Mean RT, accuracy, and proportion of "yes" responses for Experiments 1a (left) and 1b (right) are plotted as a function of cue–probe interval in Fig. 2. Contrasts evaluating distance, cue–probe interval, and their interaction are presented in Table 1. The data from both experiments show strong distance effects in RT and accuracy, suggesting that attention was focused on the target item but not sharply enough to exclude adjacent memory items, replicating Logan et al. (2021). The data from both experiments show strong cue–probe interval effects, suggesting that focusing attention on the cued item took time, replicating precuing effects in perceptual attention (Eriksen & Hoffman, 1973) and retro-cuing effects in visual short-term memory (Souza & Oberauer, 2016). The data from both experiments show little evidence that the focus becomes sharper as cue–probe interval increases. The interaction contrast testing for sharpening was only significant in the RT data in Experiment 1a, and the effect was very weak. It was convincingly null for accuracy in Experiment 1a, and convincingly null for both RT and accuracy in Experiment 1b.

The distance, cue–probe delay, and Distance × Cue–Probe Delay effects replicated across experiments for both RT and accuracy. Table 1 contains independent-samples *t* tests comparing the mean contrast values from the two experiments. None of the differences approached significance, indicating a successful replication. Focusing attention appears to depend on cue–probe interval in the same way when list-to-probe interval is fixed (1a) as when list-to-cue interval is fixed (1b).

Experiment 2: compatibility effects

The second experiment manipulated the compatibility between the cued item and the uncued items in the probe displays by manipulating the nontarget items (see Fig. 3). *Same* probes presented letters from the memory list, which were compatible for "yes" responses (list = ABCDEF, probe ABCDEF) and incompatible for "no" responses (probe ABDCEF). *Different* probes presented letters that were not from the memory list, which were incompatible for "yes" responses (probe GHCJKL) and compatible for "no" responses (GHDJKL). Thus, the compatibility effect is a crossover interaction between same versus different probes and "yes" versus "no" responses.

The compatibility effect measures resistance to distraction (Eriksen & Eriksen, 1974). Our models of the episodic flanker task propose two components that account for compatibility effects (Logan et al., 2021). The first is a *local match* that compares the probe with samples taken from the focus of attention on a single (local) item in the memory list. This sample includes information from the cued item and its immediate neighbors. The second is a *global match* that compares the entire probe with the entire memory list. Following zoom-lens (Eriksen & St. James, 1986), shrinking-spotlight (White et al., 2011), and dual-process (Cohen et al., 1992; Hübner et al., 2010) models of the flanker task, focusing attention on an item in memory might involve a transition from global matching to local matching over time. In model fits, the global match accounted for half of the compatibility effect, so the compatibility effect should decrease substantially as cue–probe delay increases. Experiment 2 tested this prediction. Alternatively, neither global nor local matches may begin before attention is oriented to the target, so the compatibility effect may not change as cue–probe delay increases.

Method

As in Experiment 1, we ran two versions of Experiment 2, 2a with list-to-probe interval fixed at 1,750 ms and 2b with list-to-cue interval fixed at 1,000 ms (see Fig. 1). These experiments were replications of two previous experiments that tested 33 subjects each on more cue-to-probe intervals (0–500 in steps of 100 ms), and fewer trials (576; one replication of the basic experimental design). The results are reported in the Supplementary Information. As in the original versions of Experiment 1, the inferential statistics were

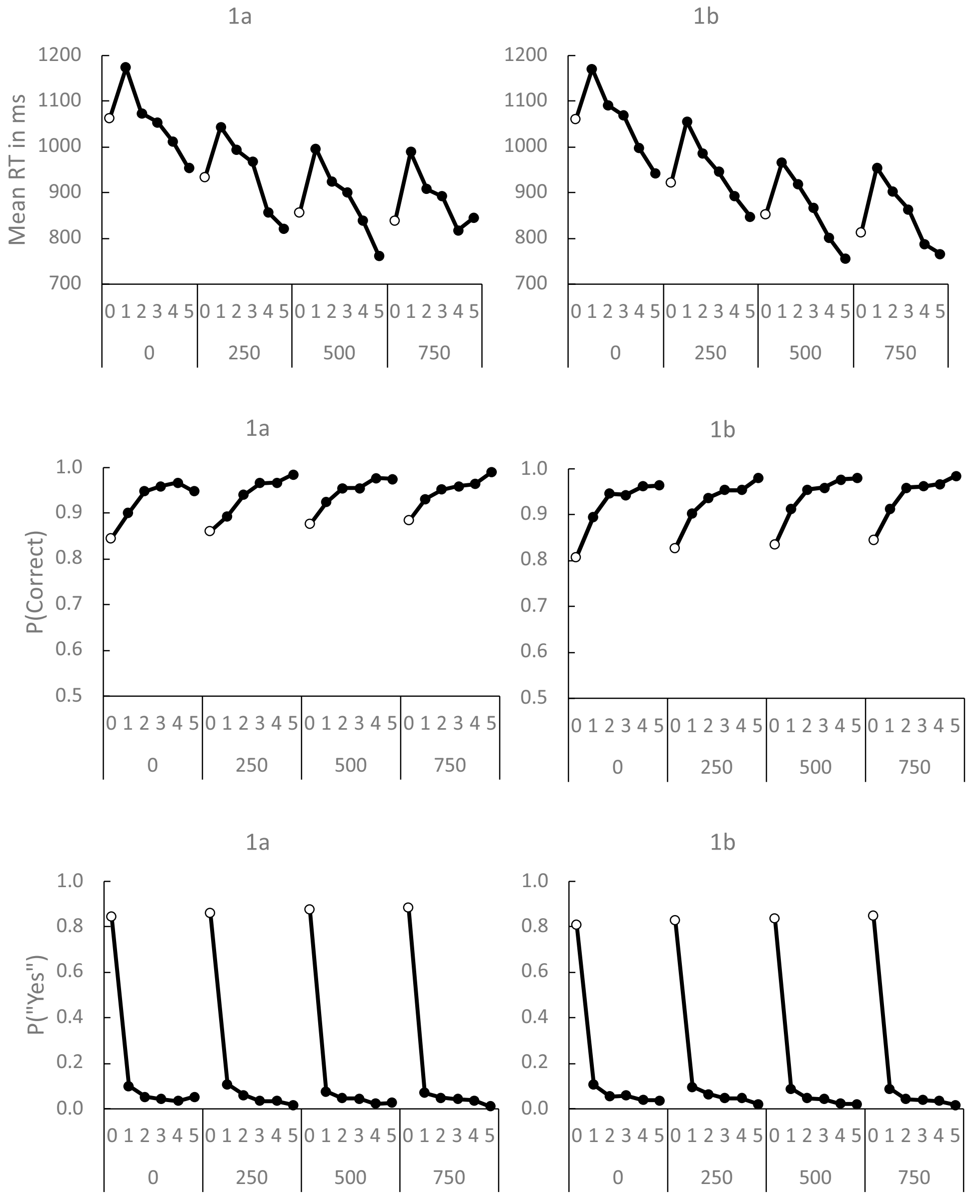

Fig. 2 Mean response time (RT; top row), accuracy (middle row), and probability of saying “yes” (bottom row) in Experiments 1a (left panels) and 1b (right panels) as a function of cue–probe interval (0–750 ms) and distance (0–5) between the position of the cued item in the probe and its position in the memory list. The positions match at distance = 0 (“yes” response; open circles) and mismatch at distances 1–5 (“no” response; filled circles). The distance contrast is defined for distances 1–5

clear and supported the same conclusions as the inferential statistics in the present experiments, but the data patterns were noisy because there were fewer observations per cell. Consequently, we replicated the experiments with fewer cue-to-probe intervals (0–750 ms in 250-ms steps) and more trials (768; two replications of the basic experimental design) to reduce experimental noise and produce more stable patterns.

Subjects We planned to run 32 subjects. Logan et al. (2021) assessed the compatibility effect in four experiments with 32 subjects each. Across experiments, 120/128 subjects showed the compatibility effect in RT and 120/128 subjects showed it in accuracy, so we were confident we could replicate the compatibility effect in samples of 32 subjects. Experiments 2a and 2b each included 32 subjects recruited online through Prolific with the same exclusion criteria as Experiments 1a and 1b. Subjects were matched on reported age (2a: M = 30.7 years, SD = 7.0 years, 1 withheld; 2b: M = 33.0 years, SD = 5.4 years) and gender (2a: 20 males, 12 females; 2b: 23

Table 1 Linear contrasts evaluating distance and cue–probe interval effects and their interaction in response time and accuracy data from Experiments 1a and 1b

Contrast	*t*(31)	*SEM*	*p*	JZS BF	*N* > 0
		Experiment 1a response time			
Distance	10.1343	49.1186	<.0001	402.060,633	30*
Interval	14.4559	45.2372	<.0001	2.45×10^{12}	32*
D × I	2.0877	293.4715	.0451	1.2666	15
		Experiment 1a P(Correct)			
Distance	6.4745	.0222	<.0001	50,489	27*
Interval	4.6794	.0190	.0001	449,8244	26*
D × I	0.4351	.0873	.6665	0.2062	15
		Experiment 1b response time			
Distance	15.1075	34.6318	<.0001	7.77×10^{12}	32*
Interval	13.3223	55.8411	<.0001	3.01×10^{11}	32*
D × I	0.1634	236.4591	.8713	0.1912	20
		Experiment 1b P(Correct)			
Distance	6.1805	.0256	<.0001	23,326	25*
Interval	3.4907	.0260	.0015	23.1844	22*
D × I	0.2548	.0828	.8006	0.1946	14
		Experiment 1a versus 1b response time			
Contrast	*t*(62)	*SEM*	*p*	JZS BF	
Distance	0.4297	35.1750	.6689	.2761	
Interval	1.2722	55.5404	.2081	.5055	
D × I	0.6903	349.0087	.4926	.3126	
		Experiment 1a versus 1b P(Correct)			
Distance	0.4281	.0255	.6701	.2761	
Interval	1.2014	.8369	.2342	.4697	
D × I	.0633	.1785	.9497	.2558	

*$p < .05$ by sign test

males, nine females). Subjects who participated in Experiments 1a or 1b were excluded from these experiments.

Apparatus and stimuli These were the same as in Experiments 1a and 1b except for the probe displays, which contained capital letters in all positions (see Fig. 3). *Same* probes presented letters from the memory list in the same positions in the probe list (e.g., ABCDEF => ABCDEF) and *different* probes presented letters in the uncued positions that had not appeared in the memory list (e.g., ABCDEF => GHCJKL). Each position was probed equally often. Half of the trials required "yes" responses (distance = 0) and half required "no" responses (distance = ±1 or ±2). Distance was manipulated as in Experiments 1a and 1b, by substituting a letter ±1 or ±2 positions away from the cued position in the memory list for the target (see Fig. 3).

Procedure The basic design required 96 trials to include all positions and same and different context conditions. Probes appeared equally often in each probe position, and there were eight trials for each probe position, four in which the probe contained the target letter and four sampling lures from positions ±1 and ±2 away from the cued position. With four cue–probe intervals, the complete design required 384 trials. We ran two replications of the complete design, resulting in 768 trials. The order of trials was randomized separately for each subject within each replication. The 768 trials were split into eight blocks of 96.

The procedure was the same as in Experiments 1a and 1b, using the same cue–probe intervals (0, 250, 500, 750 ms) and timing parameters (see Fig. 1). It differed primarily in same and different probe displays instead of the neutral probes in Experiments 1a and 1b (see Fig. 3). Because the main purpose of the experiments was to examine compatibility effects (interactions between same vs. different probes and "yes" vs. "no" responses) and their modulation with cue–probe interval, we included only distances of ±1 and ±2 in the "no" responses.

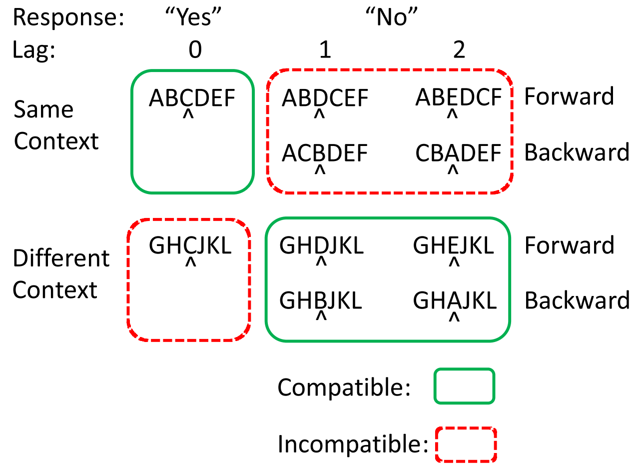

Fig. 3 Example probe displays in Experiments 2a and 2b for list ABCDEF. The left column shows Lag 0 probes that require a "yes" response. The middle and right columns show Lag 1 and Lag 2 probes, respectively, that require a "no" response. Forward probes involve lures from later positions in the list. Backward probes involve lures from earlier positions in the list. The top two rows show same context displays, in which the uncued items are the same as in the memory list. The bottom two rows show different context displays, in which the uncued items are not from the memory list. Compatible probes are circled with solid green lines. Incompatible probes are circled with dashed red lines. (Color figure online)

Data analysis In Experiments 2a and 2b, we analyzed the RT and accuracy data with four planned contrasts, calculated with the method we used in Experiments 1a and 1b, using contrast specific error terms and the same exclusion criteria. First, we analyzed compatibility effects with contrasts that evaluated the interaction between distance and context (same vs. different) using weights {2 −1 −1} for distances 0–2 in same-context probes, and weights {−2 1 1} for distances 0–2 in different-context probes. Second, we analyzed distance effects with contrasts that compared the absolute values of distance ±1 and distance ±2 with weights {1 −1}. Third, we analyzed cue–probe interval effects with linear contrast weights {3 1 −1 −3} for cue–probe intervals 0–750. Finally, we tested the predicted reduction in the compatibility effect with cue–probe interval by calculating the compatibility contrast for each cue–probe interval and testing for a reduction in the magnitude of the compatibility effect with linear contrast weights {3 1 −1 −3} for cue–probe intervals 0–750. As before, we report the number of subjects showing contrasts in the expected direction and JZS BF.

Results and discussion

Mean RT, accuracy, and the proportion of "yes" responses for Experiments 2a (left) and 2b (right) are plotted as a function of context (same vs. different), distance (0, 1, 2), and cue–probe interval in Fig. 4. Contrasts evaluating compatibility (interaction between distance and context), cue–probe interval, and the interaction between compatibility and cue–probe interval are presented in Table 2. The data from both experiments showed robust compatibility effects in RT and accuracy, replicating Logan et al. (2021)

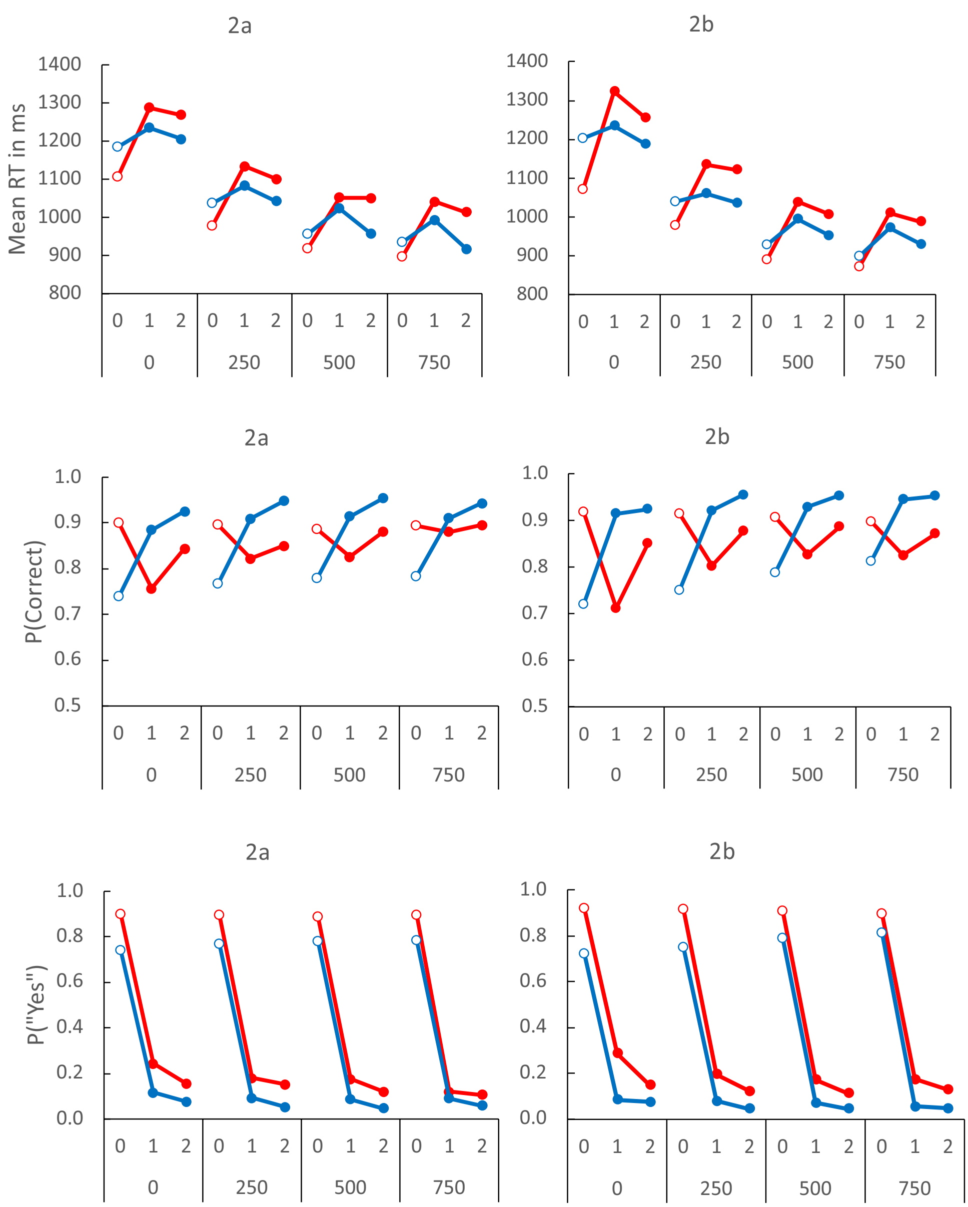

Fig. 4 Mean response time (RT; top row), accuracy (middle row), and probability of saying "yes" in Experiments 2a (left panels) and 2b (right panels) as a function of cue–probe interval, distance (0 = "yes" response; open circles; 1–2 = "no" response; filled circles), and probe context (same as memory list; different from memory list). The compatibility contrast tests for a crossover interaction between distance and context in RT and accuracy. (Color figure online)

and suggesting that neighboring items were retrieved along with the target. Both experiments showed robust cue–probe interval effects, replicating Experiments 1a and 1b and suggesting that focusing attention on the target in memory took time. There was little evidence that the compatibility effect changed with cue–probe interval. The interaction contrasts testing for a progressive reduction were convincingly null for RT and accuracy in Experiment 2b. In Experiment 2a, the RT contrast was not significant but the accuracy contrast was, suggesting a tendency for the compatibility effect to decrease with cue–probe interval. Independent samples *t* tests comparing the contrasts in the two experiments (Table 2) revealed no significant differences in any of the contrasts. Notably, the Compatibility × Interval interactions for RT and accuracy did not differ between experiments. Consequently, we conclude that the ability to resist distraction does not change as cue–probe interval increases. The preliminary versions of Experiments 2a and 2b revealed no significant interactions between compatibility and cue–probe interval in RT or accuracy, further suggesting that the interaction is not robust.

Table 2 Linear contrasts evaluating compatibility and cue–probe interval effects and their interaction in response time and accuracy data from Experiments 2a and 2b

Contrast	$t(31)$	*SEM*	*p*	JZS BF	N > 0
		Experiment 2a response time			
Compatibility	7.1522	47.4577	<.0001	293,701	28*
Interval	17.2040	45.9934	<.0001	2.49×10^{14}	32*
C × I	1.9097	176.8135	.0655	0.9434	22*
		Experiment 2a P(Correct)			
Compatibility	2.7593	.0565	.0096	4.5481	23*
Interval	5.6733	.0206	<.0001	6,117	25*
C × I	2.8429	.1644	.0078	5.4193	21
		Experiment 2b response time			
Compatibility	7.2936	50.6144	<.0001	422,266	31*
Interval	26.3383	32.7577	<.0001	3.71×10^{19}	32*
C × I	1.6177	162.7218	.1159	0.6094	17
		Experiment 2b P(Correct)			
Compatibility	3.0355	.0578	.0048	8.2100	25*
Interval	5.2873	.0261	<.0001	2,208	27*
C × I	1.0426	.1287	.3052	0.3107	18
		Experiment 2a versus 2b response time			
Contrast	$t(62)$	*SEM*	*p*	JZS BF	
Compatibility	.4286	69.3833	.6697	0.2761	
Interval	1.2665	56.4664	.2101	0.5024	
C × I	.3097	240.2944	.7578	0.2660	
		Experiment 2a versus 2b P(Correct)			
Compatibility	.2416	.0808	.8099	0.2618	
Interval	.6270	.0332	.5329	0.3017	
C × I	1.5959	.2088	.1156	0.7445	

*$p < .05$ by sign test

The bottom panel of Fig. 4 shows that the probability of saying “yes” is higher when the context is the same than when it is different. This is not a simple response bias effect because the increase is conditional on the context (same or different) rather than the required response (“yes” or “no”). We interpret it as evidence that flanking letters in the context provide input to the decision process, increasing the drift rate for “yes” responses when the context is the same and increasing the drift rate for “no” responses when the context is different, as our models predict (Logan et al., 2021).

General discussion

The experiments replicated distance and compatibility effects in the episodic flanker task (Logan et al., 2021), analogous to distance and compatibility effects in the perceptual flanker task (Eriksen & Eriksen, 1974; Eriksen & Hoffman, 1973). The similar effects suggest that the same spotlight of attention that is turned outward in the perceptual flanker task is turned inward on memory in the episodic flanker task. Memory retrieval is selective attention turned inward.

The experiments showed strong effects of cue–probe interval, reflecting the time required to orient attention to the cued location in the memory list, replicating precuing effects in perceptual attention (Eriksen & Hoffman, 1973) and retro-cuing effects in visual short-term memory (Souza & Oberauer, 2016). The main question was whether distance and compatibility effects would decrease with cue–probe interval, as if attention becomes more sharply focused as cue–probe interval increases. The results for distance were clearly null. There was no evidence of a reduction in distance effects in either RT or accuracy. The sharpness of the focus does not seem to increase with cue–probe delay, suggesting that the focus might not be set until attention is oriented to the cued item in the memory list (Smith & Ratcliff, 2009).

The results for compatibility were less clear. There was no evidence for a reduction in compatibility effects with cue–probe delay for RT and accuracy in Experiment 2b or in either of the preliminary versions of Experiment 2, as if distractors are only filtered out after attention is oriented to the item. Experiment 2a showed a significant reduction in the compatibility effect for accuracy. Considering Experiment 2b and the preliminary experiments, we conclude there was little or no change in the compatibility effect with cue–probe interval, suggesting that distractors are only filtered out after attention has focused on the cued item.

If attention does not become more sharply focused on memory and the ability to resist distraction does not improve as cue–probe delay increases, why does performance get better? We think the improvements may reflect orienting attention to the cued position in the memory list. While we have not modeled this process, it must include encoding the position of the cue and using it to search through the memory list for the cued position (Logan, 1995). The durations of the encoding and search processes should be affected by the nature of the cues and the (serial) structure of the memory list (Logan et al., 2021). If memory retrieval is attention turned inward, we may be able to understand internal orienting by applying models and methods that have illuminated investigations of perceptual encoding and visual search to memory tasks.

The time-course function allows us to measure the duration of the orienting process without specifying the computations it performs, adapting models of attention switching (Sperling & Weichselgartner, 1995) and task switching time (Logan & Bundesen, 2003). Our model defines RT_{base} as the RT when attention is oriented to the target (cue–probe interval ≈ ∞). If attention is not yet oriented to the target (cue–probe interval = 0), RT = $RT_{base} + OT$ (orienting time). We assume that OT is distributed exponentially. As cue–probe interval increases from 0 to ∞, the orienting process is more likely to have finished, so RT decreases from $RT_{base} + OT$ to RT_{base}:

$$RT = RT_{base} + \mu_{OT} \bullet \exp\left(-CPI/\mu_{OT}\right), \tag{1}$$

where CPI is cue–probe interval and μ_{OT} is the mean of the OT distribution—our measure of the duration of the orienting process.

Figure 5 plots predicted and observed time-course data from Experiments 1a, 1b, 2a, and 2b. The predictions were derived by fitting Eq. 1 to mean RTs for each cue–probe interval for each subject and plotting the average values. The averages of subjects' parameter values and measures of goodness of fit (*rmsd* = root mean squared deviation between observed and predicted values; *r* = Pearson correlation) are presented in Table 3. The fits were good.

It is interesting that the estimates of mean OT were very similar across experiments: Orienting took between 218 and 292 ms. Impressed by this similarity, we fit Eq. 1 to mean RTs (across subjects) from a retro-cuing study of visual short-term memory by Tanoue and Berryhill (2012), which measured RT at 6 cue–probe intervals ranging from 100 to 700 ms. Observed and predicted values are plotted in Fig. 5 and the best-fitting parameters and measures of goodness of fit appear in Table 3. The estimate of mean OT—244 ms—fell in the middle of the range of means from the present experiments. Fits of Eq. 1 to our original experiments (reported in Supplementary Information) also yielded values in the same range. Given the similarity of orienting times, it is tempting to think that the orienting process is performing the same computations across experiments and paradigms that engage focal attention in different ways. This temptation is encouraged further by the null effects of cue–probe delay on the sharpness of focus and resistance to distraction. They suggest that focusing on memory occurs after orienting to memory, like the eyes focus on a target after they move to it.

The episodic flanker procedure is similar to procedures in studies of visual short-term memory, which use similar list lengths, display durations, retention intervals, and retrieval tasks. Our results have implications for that literature. Our time-course results are especially relevant to retrocuing studies (Souza & Oberauer, 2016), which investigate the benefits of valid retro cues relative to invalid cues and no cues. Those studies allow us to ask about the nature of the processing engaged by valid retro cues, to identify it with attention and retrieval, and to ask how sharply it can focus on an item and how effectively it can resist distraction. Theories of retro-cue benefits do not apply directly to our results because we have no invalid-cue or no-cue conditions to calculate benefit. The theory closest to our thinking is the *retrieval head start* theory of Souza et al. (2016), which says that retro cues allow retrieval to start before the probe display is presented. Our results suggest that retro cues allow orienting to the target location in memory to start before the probe, but the decision process—retrieval itself—does not begin until the probe is presented. More generally, our results show strong effects of cue–probe interval on RT and motivate further investigations of RT in retro cuing.

Our theoretical perspective is different from typical approaches to visual short-term memory. It is inspired by theories of serial and free recall, which do not draw sharp distinctions between short-term and long-term memory and explain list length effects as interference rather than limited slots or resources (Brown et al., 2007; Farrell, 2012; Henson, 1998; Howard & Kahana, 2002; Howard et al., 2015; Lewandowsky & Farrell, 2008). These theories focus on how the structure of a list is represented in memory (noisy coding, position coding, item coding) and how that structure is used to guide retrieval of items. Our theories of the episodic flanker task assume attention to memory involves applying a retrieval cue to a memory structure—different models propose different structures. The present results suggest the retrieval cue is applied only after navigating through the structure to orient to the target. Theories of visual short-term memory

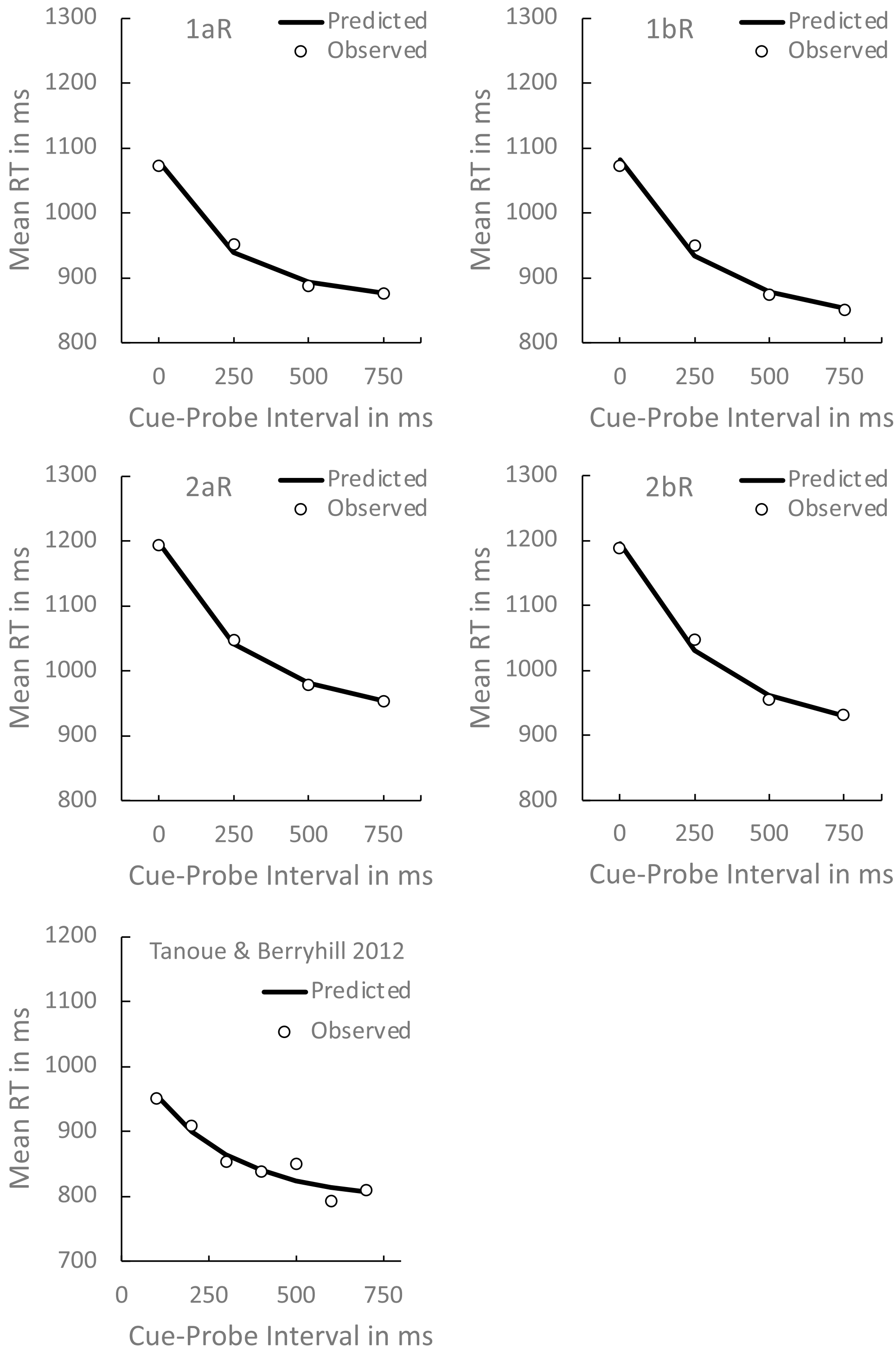

Fig. 5 Observed and predicted mean response times (RT) as a function or cue–probe interval in Experiments 1a, 1b, 2a, 2b (top and middle panels) and in the valid cue condition in Tanoue and Berryhill's (2012) Experiment 1. Predictions are from the exponential orienting model in Eq. 1

focus more on item representations (e.g., precision) than on the structures they are bound to. The episodic flanker task addresses RT and accuracy, whereas visual short-term memory and serial recall tasks focus primarily on accuracy. The benefits of considering both RT and accuracy are well known (Ratcliff, 1978). The episodic flanker task uses letters—"verbal" material—whereas studies of visual short-term memory use colors, shapes, orientations,

Table 3 Mean parameter values and measures of goodness of fit for the exponential orienting model (Eq. 1) in Experiments 1a, 1b, 2a, 2b (standard errors of the means in brackets), and in the fits to the means across subjects in valid cue condition of Tanoue and Berryhill's (2012) Experiment 1

Data source	RT_{base}	μ_{OT}	*rmsd*	*r*
Experiment 1a	834 (32)	218 (17)	16.60 (1.38)	.9465 (.0172)
Experiment 1b	828 (37)	254 (23)	17.69 (1.81)	.9504 (.0150)
Experiment 2a	928 (35)	268 (19)	15.89 (1.54)	.9732 (.0074)
Experiment 2b	904 (40)	292 (13)	19.41 (1.84)	.9766 (.0048)
T&B 2012	793	244	13.62	.9636

RT_{base} = all components of RT besides orienting; μ_{OT} = mean of the orienting time distribution

and so forth But letters and words are visual categories, apprehended by binding letters to structures that represent words to access the same semantics (Dehaene & Cohen, 2011; Grainger, 2018; Houghton, 2018; Logan, 2021). Letters, words and features may engage the same encoding and retrieval processes.

These differences are not insurmountable. There is much common ground between our thinking, theories of serial and free recall, and theories of visual short-term memory. They all assume that items are bound to some kind of structure that supports retrieval and they all account for similar—or at least related—phenomena. There is much to be learned by combining the different perspectives and exploring relations among them.

Open practices statement The data, experimental programs, analysis programs, and analysis results are available online (https://osf.io/swyp6/). None of the experiments was preregistered.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13423-022-02222-w.

Acknowledgements This research was supported in part by National Science Foundation grant BCS 2147017.

References

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*, 539–576.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523–547.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62*, 73–101.

Cohen, J. D., Servan-Schreiber, D., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. *The American Journal of Psychology*, 239–269.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences, 15*, 254–262.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods, 47*, 1–12.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*(4), 501–517.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16*(1), 143–149.

Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception & Psychophysics, 14*(1), 155–160.

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics, 40*, 225–240.

Estes, W. K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review, 104*(1), 148–169.

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review, 119*, 223–271.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16*, 129–135.

Grainger, J. (2018). Orthographic processing: A 'mid-level' vision of reading. *Quarterly Journal of Experimental Psychology, 71*, 335–359.

Henson, R. N. (1998). Short-term memory for serial order: The start–end model. *Cognitive Psychology, 36*(2), 73–137.

Houghton, G. (2018). Action and perception in literacy: A common code for spelling and reading. *Psychological Review, 125*(1), 83–116.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*(3), 269–299.

Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review, 122*, 24–53.

Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological Review, 117*(3), 759–784.

James, W. (1890). *The principles of psychology*. Holt.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review, 20*, 228–242.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 49, pp. 1–48). Academic Press.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28*, 103–174.

Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103*(4), 603–649.

Logan, G. D. (2002). An instance theory of attention and memory. *Psychological Review, 109*, 376–400.
Logan, G. D. (2005). The time it takes to switch attention. *Psychonomic Bulletin & Review, 12*, 647–653.
Logan, G. D. (2021). Serial order in perception, cognition, and action. *Psychological Review, 128*, 1–44.
Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit taskcuing procedure? *Journal of Experimental Psychology: Human Perception and Performance, 29*, 575–599.
Logan, G. D., Cox, G. E., Annis, J., & Lindsey, D. R. B. (2021). The episodic flanker effect: Memory retrieval as attention turned inward. *Psychological Review, 128*, 397–445.
Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 51–108.
Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychological Review, 116*, 283–317.
Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception & Psychophysics, 78*, 1839–1860.
Souza, A. S., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory: Retro-cues enhance retrieval and protect from visual interference. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 890–910.
Sperling, G., & Weichselgartner, E. (1995). Episodic theory of the dynamics of spatial attention. *Psychological Review, 102*(3), 503–532.
Tanoue, R. T., & Berryhill, M. E. (2012). The mental wormhole: Internal attention shifts without regard for distance. *Attention, Perception, & Psychophysics, 74*, 1199–1215.
White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive Psychology, 63*(4), 210–238.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.