

Reviewed Preprint
v1 • August 7, 2026
Not revised

Characterising semantic prioritisation in visual working memory

✉ For correspondence:

kerren@cbs.mpg.de

lindedomingo@ugr.es

Competing interests: No competing interests declared

Funding: See page 24

Reviewing editor: Roberto Bottini, University of Trento, Italy

© 2026, Kerrén et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Casper Kerrén¹ ✉, Carlos González-García^{2,3}, Juan Linde-Domingo^{2,3} ✉

¹Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany • ²Mind, Brain and Behavior Research Center, University of Granada, Granada, Spain • ³Department of Experimental Psychology, University of Granada, Granada, Spain

eLife Assessment

This **useful** study addresses a timely question about semantic prioritisation in visual working memory, using behavioural manipulations and drift-diffusion modelling. However, the strength of evidence is **incomplete** for the broader claims about working-memory representations because the main interpretation relies on indirect inferences from non-decision time, which cannot uniquely identify memory access or retrieval.

<https://doi.org/10.7554/eLife.111859.1.sa3>

Abstract

Cognitive operations require recently encountered information to remain available beyond the moment of sensory input. However, how such transient representations are accessed, and how they differ from sensory processing and long-term memory, remains unclear. Here, we combine hierarchical drift-diffusion modelling with behavioural manipulations to dissociate the decision processes underlying feature prioritisation in visual working memory. We first reanalysed a previously collected working memory dataset to characterise semantic and perceptual judgements at the level of latent decision processes. Semantic judgements were associated with reduced non-decision time across conditions, indicating faster access to task-relevant information, while advantages in evidence accumulation emerged selectively under higher cognitive demands. Two further experiments manipulated attentional prioritisation using retro-cues and dissociated the effects of interference from mere maintenance. Across manipulations, semantic prioritisation was selectively expressed in pre-accumulation processes and was amplified when representations fell outside the focus of attention or had to be maintained under interference. Together, these results suggest that semantic representations remain more readily accessible than perceptual details when working memory representations fall outside the focus of attention, consistent with a shift towards more abstract, long-term memory-like formats under conditions of limited attentional support.

Introduction

Adaptive behaviour depends on the ability to use information that is no longer available by current sensory input. To achieve this, cognitive systems must therefore sustain internal representations over short intervals and make them available for ongoing computations. This capacity, known as working memory (WM ^{1–3}), enables the use of task-relevant information beyond the moment of perception. Despite its centrality, it remains unclear what format information takes in WM, how it is accessed, and to what extent WM representations resemble those instantiated during perception or those reconstructed from long-term memory (LTM).

A large body of research has investigated WM using delay-based paradigms, often involving the maintenance of a single stimulus or feature ^{3,4}. These studies have demonstrated that information held in WM can be decoded from brain sensory cortical regions typically associated with perceptual processing ^{5,6}, motivating influential sensory recruitment accounts. From this

perspective, WM maintenance involves the continued engagement of perceptual representational formats. At the same time, converging behavioural and neural evidence indicates that WM representations are frequently transformed relative to their perceptual origin ^{7–9}. Rather than preserving a veridical sensory trace, WM may prioritise task-relevant dimensions, abstracted and semantic stimulus features ^{10,11}, or anticipated actions, depending on behavioural and attentional demands ^{12–14}.

Beyond the maintenance of a single item, WM is generally assumed to support the concurrent retention of multiple representations ^{15–17}. However, only some of these can occupy the focus of attention at any given moment ^{17–19}. How information outside this focus is retained and accessed remains a matter of debate. Contemporary proposals range from transient synaptic traces that persist without sustained neural firing ^{20,21}, to distributed storage across cortical hierarchies at varying levels of abstraction. A more classical view holds that unattended WM contents may rely on mechanisms closely related to those supporting long-term memory, particularly when they must be reactivated after a period of reduced attentional priority ^{22,23}. Recent theoretical and empirical work has extended this idea, suggesting that activity-silent WM may share computational properties with episodic memory retrieval ^{23,24}. From this perspective, whether information remains within the focus of attention, rather than how long it has been stored, may be a critical determinant of how it is accessed and used.

Insights from the episodic memory literature highlight a critical distinction between how information is accessed from memory and how it is processed during perception. During visual perception, neural processing typically unfolds along a feedforward trajectory in which low-level sensory attributes are generally resolved before higher-level, conceptual properties ^{25–29}. In contrast, during episodic retrieval, conceptual or semantic information often becomes available more rapidly than fine-grained perceptual detail ^{30,31}. This inversion of temporal dynamics has been interpreted as evidence that episodic retrieval is reconstructive, potentially initiated by pattern completion mechanisms that reinstate cortical representations in a top-down manner ^{32–34}. Comparable temporal reversal patterns have also been observed in visual imagery ^{35,36}, reinforcing the idea that internally generated representations differ fundamentally from perceptually driven ones.

We previously asked whether similar access dynamics characterise WM. In ³⁷, we reported a robust prioritisation of semantic over perceptual information when multiple items were maintained in working memory. Participants were faster, and often more accurate, in reporting the semantic as compared to the perceptual format. However, in that study temporal distance was necessarily confounded with attentional allocation, leaving open whether lag itself, or the likelihood of attentional disengagement, drove the effect. These response dynamics resembled those observed during episodic memory retrieval ^{30,31} and contrasted with those typically found in visual perception, where lower-level visual features are accessed earlier than higher-level conceptual attributes.

Despite this evidence, the basis of semantic prioritisation in working memory remains unclear. Much of the existing debate regarding the format of working memory representations has focused on measures of accuracy or neural decodability ^{5,38}, which do not distinguish whether differences arise from representational fidelity or from the dynamics with which stored information can be accessed ³⁹. One possibility is that semantic bias reflects facilitated access to stored representations when information must be reactivated from an unattended state. Alternatively, it may arise from selection processes or decision-level biases, for example if participants pre-activate a single item following a cue or adopt different response strategies for semantic and perceptual judgements ⁴⁰. More generally, it remains unknown whether semantic prioritisation is driven primarily by attentional disruption during maintenance, by the need to reinstate information from an unattended state, or by the passage of time itself. Recent work further shows that retrospective attention can preferentially enhance perceptual information in working memory when prioritisation operates through attentional selection, suggesting that different internal processes may also influence distinct stages of information access and decision formation ¹⁴.

In the present study, we build directly on these earlier findings to examine the conditions under which semantic prioritisation emerges in working memory. In particular, here we tested whether such bias in working memory arises prior to evidence accumulation or during the decision process itself. To adjudicate between access and accumulation-based accounts of semantic prioritisation, we applied drift diffusion modelling (DDM ^{41–43}) to dissociate pre-decisional access from evidence accumulation. First, we reanalysed our previous results ³⁷ using DDM. Building on these results, in Experiment 1, we combined DDM with a retro-cue manipulation to assess whether this semantic preference is amplified when attentional selection is prevented and multiple items must remain concurrently maintained ⁴⁰. In Experiment 2, we independently manipulated interference and delay during the maintenance interval to dissociate the effects of attentional disruption and time on the accessibility of semantic and perceptual information ^{44,45}. Together, our results show that semantic prioritisation in working memory is a robust phenomenon that potentially reflects faster access to stored representations, whereas accumulation advantages are conditional on attentional state.

Results

Drift-diffusion modelling of Kerrén et al (2022)

In ³⁷, we asked whether the semantic prioritisation observed during episodic memory retrieval (e.g., ³¹) also characterises visual working memory (VWM). Across two experiments, we observed a robust semantic advantage when participants were required to maintain more than one item simultaneously. This semantic prioritisation was systematically modulated by both memory load (the number of items studied, ranging from one to four) and the temporal interval between study and test (hereafter referred to as lag).

Importantly, this prioritisation of semantic over perceptual information is not a trivial consequence of stimulus properties (Fig. 1A ³⁷). In a previous publication, using the same material and feature space, perceptual information (line-drawing vs. photograph) was accessed earlier than semantic features (animacy) when judgements were made during visual processing, without a memory demand ³¹. This reversal suggests that the relative accessibility of perceptual and semantic features depends on whether processing is perceptually driven or internally generated.

Here, we reanalysed data from ³⁷ (Experiment 2), in which participants ($n = 103$ after exclusions) performed a VWM task requiring maintenance of one to four object images across a short delay (Fig. 1B ³⁷). At the test, participants answered either a semantic or a perceptual question about one of the studied items. Despite the robust behavioural effects observed in this dataset, it remains unclear what drives the semantic bias. More specifically, we hypothesized that semantic prioritisation could reflect (i) more optimal reactivation of semantic features during retrieval, (ii) selection processes operating on equally reactivated representations, or (iii) a combination of both mechanisms.

To address these questions, we applied hierarchical DDM to dissociate pre-decisional access from evidence accumulation processes. Probe type (semantic versus perceptual) was allowed to modulate non-decision time (t), indexing processes occurring prior to evidence accumulation (e.g., probe processing and preparatory processes preceding decision formation), drift rate (v), indexing the efficiency with which evidence contributes to accumulation, or both parameters jointly. Under this framework, differences in non-decision time are often interpreted as reflecting differences in processes occurring prior to evidence accumulation, whereas differences in drift rate reflect differences in the efficiency with which available information contributes to the decision process. Where relevant, we discuss interpretations of these effects in relation to retrieval or access processes. All analyses were conducted separately for each memory load and lag, allowing us to assess how access-related and accumulation-related dynamics varied as a function of attentional and temporal demands.

When trials were grouped by memory load, a highly consistent pattern emerged for non-decision time. Semantic judgements were associated with reliably shorter non-decision times than perceptual judgements at every load. The posterior mean difference in non-decision time

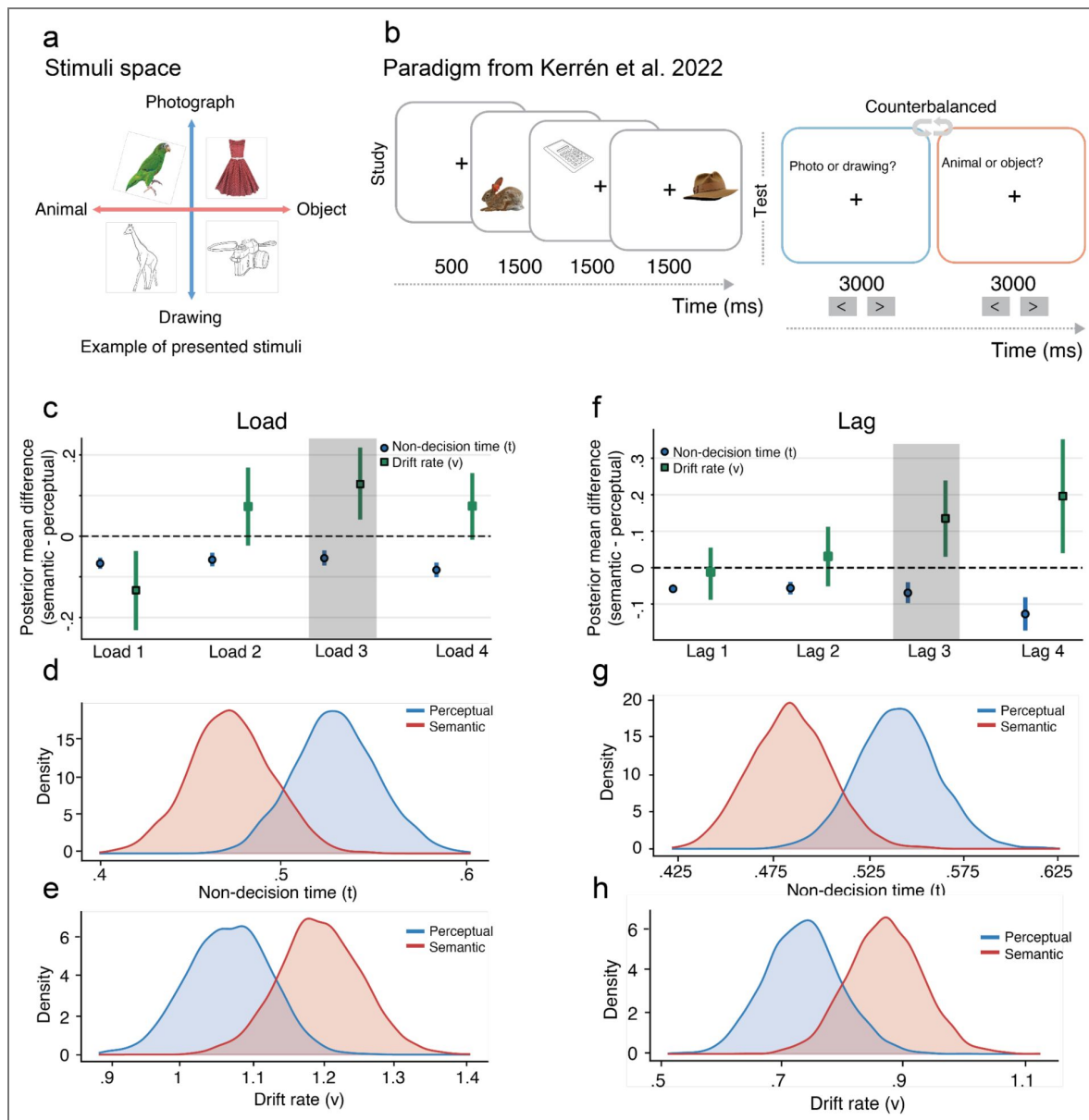


Figure 1. Paradigm and results of reanalysis of Kerrén et al.

a The stimuli space for all experiments followed the same rationale. Images varied along a semantic (Animal / Object) and a perceptual (Photograph / Drawing) dimension. **b** The paradigm used in ³⁷, where participants studied between 1 and 4 items and were later tested on one of the items using either a perceptual or semantic probe. **c** Using drift-diffusion modelling, we found reliable evidence for non-decision time (NDT) and drift rate differences between semantic and perceptual questions for load 3 (highlighted with grey bar) **d** Zooming in on NDT in load 3, semantic probes had lower NDT as compared to perceptual ones. **e** The effect was flipped for drift rate, where we observed a higher drift rate for semantic probes. **f** When instead organising the data according to lag between study and test, we observed a continuous build-up of increasing difference in drift rate between semantic and perceptual probes, with reliable evidence in lag 3 and 4. NDT was significant for all lag levels **g** In lag 3, we found very similar results to load 3, with a shorter NDT for semantic questions. **h** There was a smaller, yet reliable, difference in drift rate between semantic and perceptual questions.

(semantic minus perceptual) was negative for loads 1 through 4 (load 1: $-.067$ s, 94% HDI $[-.080, -.053]$; load 2: $-.058$ s, $[-.074, -.041]$; load 3: $-.054$ s, $[-.072, -.035]$; load 4: $-.083$ s, $[-.101, -.065]$), and in all cases the 94% highest density interval excluded zero. Thus, semantic information was accessed more rapidly than perceptual information at pre-decisional stages across all memory loads, indicating an advantage in processes occurring prior to evidence accumulation that was not contingent on the number of maintained items.

In contrast, effects of memory load on drift rate revealed a load-dependent reversal. When only a single item was maintained (load 1), semantic judgements showed reliably lower drift rates than perceptual judgements (mean difference = $-.133$, 94% HDI $[-.231, -.037]$), indicating slower evidence accumulation for semantic decisions under conditions of focused attention. At higher loads, however, the direction of the effect shifted. For load 2, the posterior mean difference was positive (0.073), but the highest density interval overlapped zero ($[-.023, .169]$), providing inconclusive evidence. For load 3, semantic judgements exhibited a reliably higher drift rate than perceptual judgements (mean difference = $.128$, 94% HDI $[.041, .218]$), indicating higher evidence accumulation efficiency when multiple items had to be maintained concurrently. For load 4, the posterior mean difference remained positive ($.074$), but the highest density interval again overlapped zero ($[-.009, .155]$). Together, these results indicate that while semantic judgements show a robust advantage in processes occurring prior to evidence accumulation across loads, differences in evidence accumulation efficiency emerge selectively under higher mnemonic demands, particularly when attentional resources must be distributed across multiple items.

When trials were instead grouped by lag between study and test, we tested the prediction that semantic information would be accessed more rapidly at pre-decisional stages across all lags, but that advantages in evidence accumulation would emerge preferentially at longer delays. Mimicking the previous results, semantic judgements were associated with reliably shorter non-decision times than perceptual judgements at every lag. The posterior mean difference in non-decision time (semantic minus perceptual) was negative for all lags (lag 1: $-.058$ s, 94% HDI $[-.068, -.048]$; lag 2: $-.056$ s, $[-.073, -.039]$; lag 3: $-.069$ s, $[-.097, -.040]$; lag 4: $-.127$ s, $[-.172, -.081]$), and in all cases the 94% highest density interval excluded zero. This indicates robust evidence that semantic judgements were associated with shorter non-decision times than perceptual judgements across all lags, consistent with an advantage in processes occurring prior to evidence accumulation, independent of the delay between study and test.

In contrast, effects of lag on drift rate revealed a graded pattern. At shorter lags (lag 1 and lag 2), the posterior mean difference in drift rate between semantic and perceptual judgements was small and uncertain (lag 1: mean = $-.012$, 94% HDI $[-.088, .055]$; lag 2: mean = $.031$, $[-.051, .112]$), with highest density intervals overlapping zero. At longer lags, however, semantic judgements showed reliably higher drift rates than perceptual judgements. For lag 3, the posterior mean difference was $.135$ (94% HDI $[.030, .239]$), and for lag 4 it increased further to $.196$ (94% HDI $[.040, .352]$), with the highest density interval excluding zero in both cases. Thus, while semantic information shows a consistent access advantage across delays, advantages in evidence accumulation emerge selectively at longer lags, suggesting that semantic information contributes more efficiently to decision formation as the temporal distance from encoding increases.

Together, these results indicate that semantic prioritisation in working memory is expressed through distinct decision components. Reduced non-decision time for semantic relative to perceptual judgements was present across all memory loads and lags, indicating a consistent advantage in processes occurring prior to evidence accumulation. In contrast, advantages in evidence accumulation emerged selectively under higher memory load, most clearly at load 3 and for larger time lags, indicating faster accumulation of evidence for the semantic features as well as faster access to these. This dissociation suggests that semantic prioritisation primarily reflects differences in processes occurring prior to evidence accumulation, which may be consistent with differences in retrieval or access-related processes, with additional contributions from accumulation efficiency when attentional demands are high.

Experiment 1

The drift-diffusion modelling results obtained from re-analysing our previous dataset suggest that semantic prioritisation in working memory arises from at least two separable components: (i) a robust advantage in pre-decisional access to semantic information, reflected in shorter non-decision times across loads and lags, and (ii) a more selective advantage in evidence accumulation that emerges under higher cognitive demands. However, these analyses alone cannot determine why semantic information is prioritised. In particular, it remains unclear whether the observed semantic advantage reflects a genuine retrieval-related benefit, where semantic representations are more readily reactivated from working memory, or whether it instead arises from a selection bias operating after information has already been accessed.

Experiment 1 was designed and pre-registered (<https://researchbox.org/5966>) to adjudicate between these possibilities by manipulating the extent to which retrieval demands were required at the time of the decision (Fig. 2a). By introducing a retro-cue that either specified the relevant memory item in advance or remained neutral, we aimed to dissociate retrieval-related processes from selection at the decision stage. In trials where a valid retro-cue is presented prior to the probe, participants are able to prioritise and access the relevant mnemonic content in advance, thereby reducing retrieval demands at the time of decision and rendering any behavioural differences between question types more consistent with feature-selection processes. In contrast, in neutral trials where the retro-cue is uninformative, retrieval and selection demands must both be resolved at probes, such that any observed semantic advantage could reflect differences in either or both processes. If semantic prioritisation primarily reflects differences in retrieval, then reducing selection uncertainty and allowing advance prioritisation of the relevant item via a valid cue should attenuate or eliminate the semantic advantage. Conversely, if semantic prioritisation reflects selection biases operating on already retrieved information, then the advantage should persist even when retrieval demands are minimised. In the following experiment, we combined behavioural measures with drift-diffusion modelling to examine how semantic and perceptual information are differentially accessed and accumulated under valid and neutral cue-conditions.

Eighty-three participants were exposed to 3 items during study, sequentially presented in three pseudorandom locations on the screen (each for 1500ms). We selected three items because this condition yielded the most reliable differences in non-decision time and drift rate in the reanalysis of our previous dataset. The task for participants was to remember the items and the locations. Immediately after the third item, a retro-cue was shown. In the valid condition, a circle was shown in the location where the questions about the item was going to be shown. In the neutral condition, three circles were shown in the location of the three studied items. Following a 1000ms delay, a question appeared on screen (“Animal or Object?” or “Photo or Drawing?”, counterbalanced; see Methods) for three seconds. Participants were instructed to push the right or left button to register a response. Once pushed, the font colour of the question changed to either green (correct) or red (incorrect). Thereafter the next question appeared. If the first question probed the semantic dimension, e.g., “Animal or Object?”, the second question probed the perceptual features, e.g., “Photo or Drawing?” (counter-balanced; see Methods). We analysed accuracy and reaction times (RTs) to the first question of each trial.

Participants were on average good at performing the task (accuracy: semantic question mean (M) = 86.57%, standard deviation (SD) = 11.86%; perceptual question M = 82.50%, SD = 9.73%) and were on average better in the semantic as compared to the perceptual question ($t(1,82) = 4.51$, $p < .01$, Bayesian t-test against zero comparing semantic and perceptual probes: $BF_{10} = 829.97$; Fig. 2b). They were also faster at responding to the semantic as compared to the perceptual questions (semantic: $M = 846$ ms, $SD = 172$ ms; perceptual: $M = 1001$ ms, $SD = 190$ ms; $t(1,82) = -9.97$, $p < .01$, $BF_{10} > 1000$; Fig. 2c). These results very much mimic previous results and suggest that semantic features are overall easier and faster to remember.

Our main goal of Exp. 1 was to understand how much the semantic prioritisation was due to a selection bias as compared to genuine memory activation. In the valid condition, the retro-cue signalled which of the three items was going to be probed, thereby allowing participants to pre-

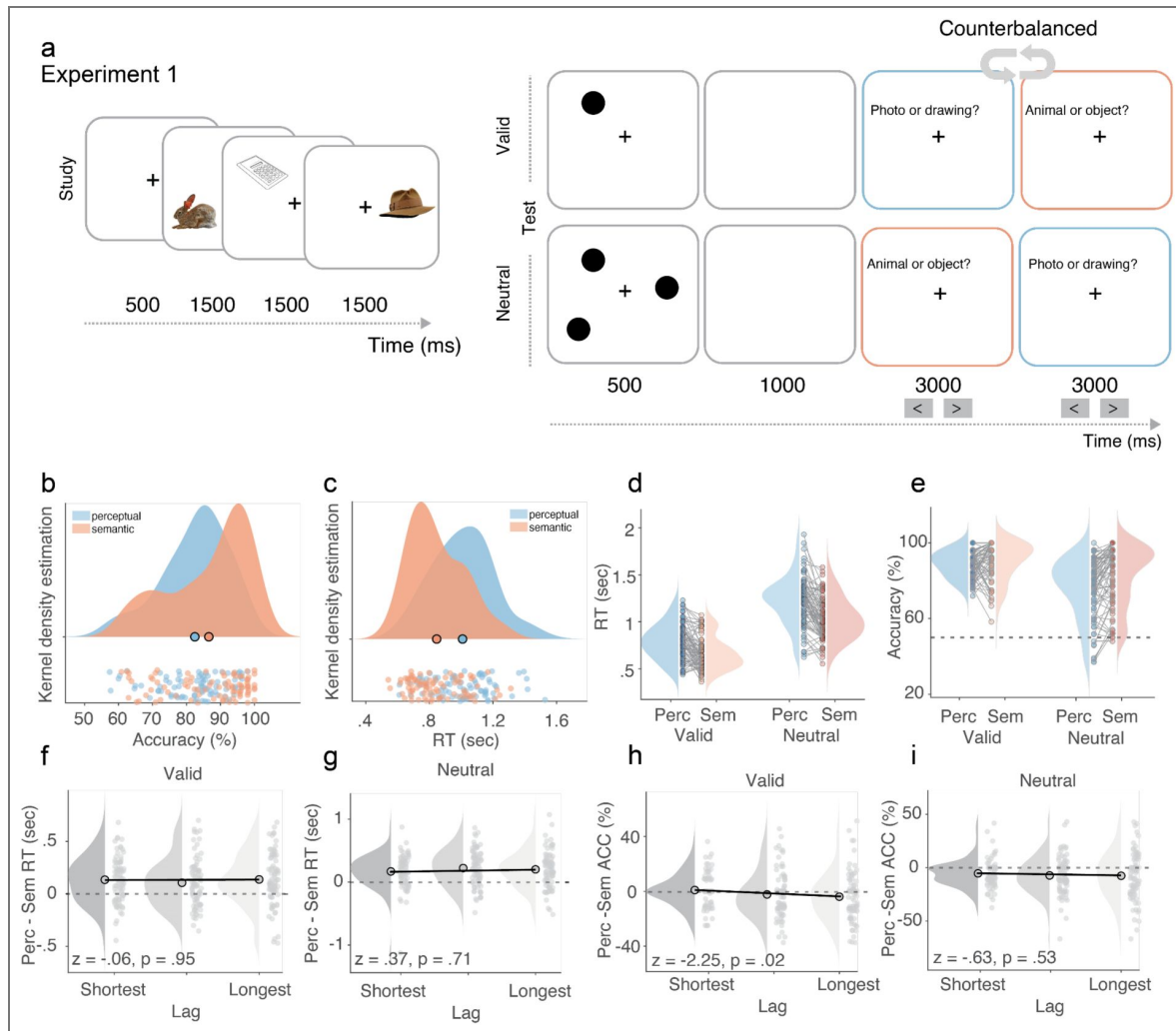


Figure 2. Paradigm and results of Experiment 1.

a Behavioural paradigm. During study, participants saw three items sequentially in pseudorandom positions around the fixation cross. After encoding, they either saw a single circle (valid condition), indicating which item was going to be tested later, or three circles (neutral condition), indicating the three positions in which items were presented during study. Following a 1000ms blank screen, they were tested for the cued item (in the valid condition), or a random item (neutral condition). **b** Average accuracy for semantic trials was higher than perceptual trials across conditions. **c** Average reaction times (RTs) for semantic trials were faster than perceptual trials. **d** We observed faster semantic RTs in both valid and neutral conditions. **e** We observed higher accuracy for the semantic questions in the neutral condition, but not in the valid one. **f** No difference between conditions for valid trials. **g** And not for neutral trials either. **h** A significant negative slope was found for the valid condition, indicating that semantic prioritisation increased with lag. **i** This was not the case in neutral trials.

activate it. In the neutral condition, such pre-activation was not possible; instead, participants had to keep all three items simultaneously in mind. Our main hypothesis was that there would be more semantic prioritisation in the neutral condition, where memory retrieval is required during the probe, as compared to the valid condition. Indeed, analysing RTs using a Bayesian t-test against zero, we observed moderate evidence ($BF_{10} = 6.20$; Fig. 2d) for more semantic prioritisation (larger difference between RTs for semantic and perceptual question) for the neutral as compared to the valid condition. Although our results suggest that selection had an impact on the format of working memory representations, we still found a reliable semantic prioritisation in both conditions, separately (valid: mean difference semantic and perceptual RTs = 128ms, SD = 170ms, $BF_{10} > 1000$; neutral: mean difference semantic and perceptual RTs = 198ms, SD = 200ms, $BF_{10} > 1000$), thus indicating that memory reactivation of previously studied items adhere to earlier availability of the semantic features.

Also, for accuracy did we observe that semantic features were more favoured in the neutral condition (difference semantic and perceptual accuracy in valid condition $M = 1.6\%$, SD = 10%; neutral condition $M = 6.64\%$, SD = 10.43%, $BF_{10} = 93.70$; Fig. 2e). Interestingly, when analysing each condition separately, we did not find any evidence for a semantic prioritisation in the valid condition (semantic $M = 90.3\%$, SD = 10%; perceptual $M = 88.75\%$, SD = 6.88%, $BF_{10} = .31$). In contrast, we did observe it in the neutral condition (semantic $M = 82.90\%$, SD = 15.97%; perceptual $M = 76.26\%$, SD = 15.68%, $BF_{10} > 1000$).

To further understand the semantic prioritisation, we split the data into lag between study and test. As the items were presented sequentially during study, the lastly presented item during study would be regarded the closest in time to test, whereas the firstly presented item during study would be the one farthest away. In our previous study³⁷ we found that when there was short time between study and test, we observed less semantic prioritisation as compared to when the lag was larger. We did not find any evidence for a positive slope in any of the conditions when assessing RTs (valid difference distance 1: $M = 136\text{ms}$, SD = 215ms; distance 2: $M = 107\text{ms}$, SD = 229ms; distance 3: 138ms, SD = 247ms, $z = -.06$, $p = .95$; neutral difference distance 1: $M = 170\text{ms}$, SD = 270ms; distance 2: $M = 224\text{ms}$, SD = 280ms; distance 3: 202ms, SD = 288ms, $z = .37$, $p = .71$; Fig. 2f and 2g). However, when assessing accuracy, we did find a negative slope for the valid condition (valid difference distance 1: $M = .95\%$, SD = 11.81%; distance 2: $M = -2.27\%$, SD = 16.91%; distance 3: -3.91%, SD = 17.67%, $z = -2.25$, $p = .02$; neutral difference distance 1: $M = -5.36\%$, SD = 12.75%; distance 2: $M = -7.29\%$, SD = 17.33%; distance 3: -7.60%, SD = 21.87%, $z = -.63$, $p = .53$; Fig. 2h and 2i). Interestingly, we found numerically higher accuracy for perceptual probes when study and test were closest in time in the valid condition, however, with moderate evidence for the null-hypothesis ($BF_{10} = .16$). Together, these results suggest that lag-related modulation of semantic prioritisation depends on the ability to focus attention on a single item, with time effects emerging only under valid cueing and not when attention must remain distributed across items.

Balanced integration score

To take into account the speed-accuracy trade-off, we used the balanced integration score (BIS), which quantifies performance in terms of both RT and accuracy with equal weight⁴⁷. Here we calculated the difference in the BIS between perceptual and semantic probes where larger values reflect more semantic prioritisation. As predicted, we found strong evidence for the more semantic prioritisation in neutral condition compared to the valid condition (valid: $M = .94$, SD = 1.46; neutral: $M = 1.89$, SD = 1.86, both $BF_{10} > 1000$ when tested against zero and $BF_{10} > 1000$ when tested for difference between valid and neutral; Fig. 3a). For valid trials, this result was likely driven by the faster RTs to semantic as compared to the perceptual probes.

When evaluating distances, we found a significant positive slope for the valid condition ($z = 2.01$, $p = .04$; Fig. 3b), suggesting that when participants could focus on one item, the semantic prioritisation was reduced, but then increased with lag. However, separate tests still revealed significant semantic prioritisation for each distance (distance 1: $M = .73$, SD = 1.73, $t(1,82) = 3.86$, $p < .01$; $BF_{10} = 94.50$; distance 2: $M = .88$, SD = 2.26, $t(1,82) = 3.56$, $p < .01$; $BF_{10} = 37.37$; distance 3: $M = 1.24$, SD = 2.56, $t(1,82) = 4.43$, $p < .01$; $BF_{10} = 613.72$). We did not find a significant slope for the

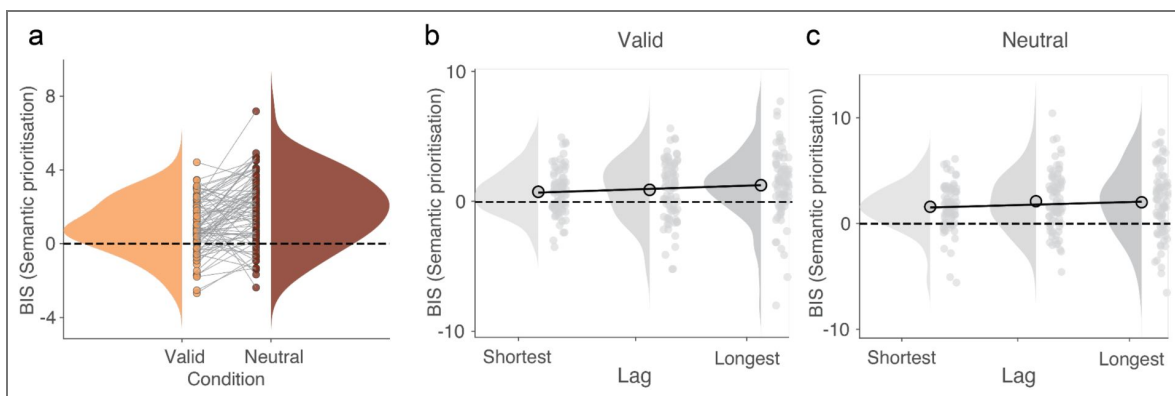


Figure 3. Balanced integration score.

a BIS was significantly higher in neutral trials (brown) as compared to valid trials (orange). **b** When organising the data in the valid condition according to lag between study and test, we observed a significant positive slope across lags, reflecting more semantic prioritisation for larger lags **c** This was not the case for the neutral condition, where a similar semantic prioritisation was present in all lag levels.

neutral condition ($z = 1.53$, $p = .13$; Fig. 3c [↗](#)), suggesting a similar semantic prioritisation for all distance levels (distance 1: $M = 1.58$, $SD = 2.08$, $t(1,82) = 6.97$, $p < .01$; $BF_{10} > 1000$; distance 2: $M = 2.10$, $SD = 2.74$, $t(1,82) = 6.99$, $p < .01$; $BF_{10} > 1000$; distance 3: $M = 2.01$, $SD = 3.13$, $t(1,82) = 5.84$, $p < .01$; $BF_{10} > 1000$).

Taken together, Exp.1 shows that the observed semantic prioritisation in VMW is mainly a genuine memory reactivation effect. However, allowing participants to pre-activate one of the three items from study has a significant effect, but it does not eliminate the semantic prioritisation completely.

Drift-diffusion modelling of Experiment 1

Together, the behavioural results of Experiment 1 indicate that semantic prioritisation in working memory is robust, but also sensitive to task demands. While valid retro-cues reduced the magnitude of the semantic advantage, consistent with a contribution of selection processes, semantic judgements were still associated with faster responses (but not higher accuracy) than perceptual judgements even when retrieval demands were minimised. This pattern suggests that semantic prioritisation cannot be explained solely by post-retrieval selection biases, but instead reflects differences in how semantic and perceptual information contribute to decision-making under different task demands.

Reaction time and accuracy alone do not allow us to determine which stages of the decision process are affected by cueing, nor whether the observed effects arise from changes in evidence accumulation, pre-decisional access, or both. To address these questions, we next applied hierarchical drift-diffusion modelling, allowing us to dissociate cue- and feature-related effects on non-decision time and drift rate, and to test whether the reduced semantic prioritisation under valid cueing reflects changes in processes occurring prior to evidence accumulation, in evidence accumulation itself, or their interaction.

To dissociate retrieval- and decision-related contributions to semantic prioritisation, we fitted a series of hierarchical drift diffusion models that allowed feature type (semantic vs perceptual) and cue validity (valid vs neutral) to modulate non-decision time (t) and drift rate (v). Model comparison indicated that a most complex model, which allowed three parameters to vary (non-decision time, drift rate and threshold), provided the best overall fit to the data. However, model comparison using PSIS-LOO-CV showed a modest difference in predictive accuracy ($\text{elpd}_{\text{loo}} = 24.26 \pm 15.45$), suggesting no reliable advantage of this more complex model over the simpler model with just v and t . Thus, we report the less complex model in the main text for parsimony, and include a comparison of models in the Supplementary Materials as exploratory.

If non-decision time reflects processes occurring prior to evidence accumulation that are influenced by retrieval demands, we would expect a larger difference in non-decision time between semantic and perceptual features in the neutral as compared to the valid condition, where retrieval demands at the time of decision are reduced. By contrast, if drift rate captures selection processes operating on already retrieved information, a feature by cue-validity interaction is also expected. However, in this case, drift-rate differences between semantic and perceptual features should be larger for valid cues than for neutral trials.

As expected, for non-decision time, we observed a reliable feature by validity interaction indicating that cueing reduced the semantic advantage in non-decision time (Fig. 4a [↗](#)). In neutral trials, semantic judgements showed substantially shorter non-decision times than perceptual judgements (semantic minus perceptual = $-.086$ s, 94% highest density interval $[-.098, -.070]$, posterior probability $< 0 = 1$), consistent with faster pre-decisional access to semantic information. In valid trials, this semantic advantage was largely eliminated ($-.005$ s, 94% highest density interval $[-.016, .002]$, posterior probability $< 0 = .858$). Correspondingly, the interaction term was positive and credibly different from zero (change from neutral to valid = $.082$ s, 94% highest density interval $[.060, .092]$, posterior probability $> 0 = 1$), indicating that the semantic minus perceptual difference became markedly less negative under valid cueing. This pattern suggests that when retrieval demands are reduced by a valid retro-cue, the semantic advantage observed in pre-accumulation processes is strongly attenuated.

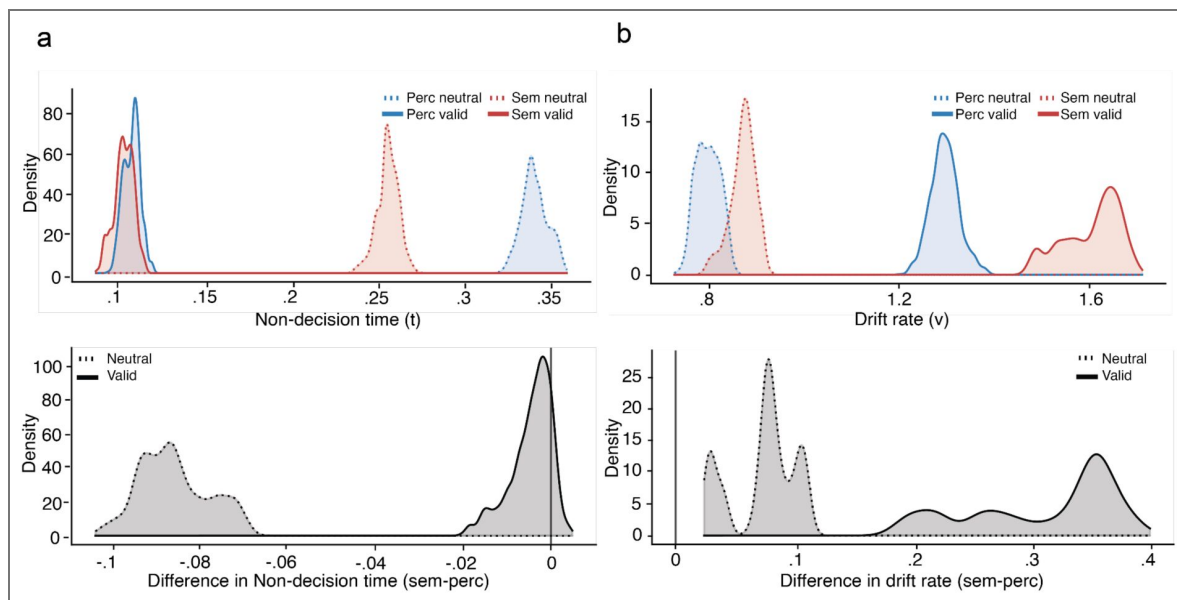


Figure 4. Drift-diffusion modelling of Experiment 1.

a (top) We observed similar non-decision time for perceptual (blue) and semantic (red) in the valid condition (solid lines), whereas for the neutral condition, perceptual trials had significantly longer non-decision time. **(bottom)** Same as top but subtracted perceptual from semantic for valid (solid line) and neutral (dashed line) conditions, separately. **b (top)** In the valid condition, we observed significantly lower drift rate for perceptual trials as compared to semantic trials. This was not the case for the neutral condition. **(bottom)** Same as top, but subtracted perceptual from semantic, for valid and neutral trials, separately.

For drift rate, we also observed an expected robust feature by validity interaction, but in the opposite direction (Fig. 4b). In neutral trials, semantic judgements already showed higher drift rates than perceptual judgements (semantic minus perceptual = .073, 94% highest density interval [.026, .108], posterior probability > 0 = 1). Critically, this semantic drift advantage was substantially larger in valid trials (.309, 94% highest density interval [.192, .378], posterior probability > 0 = 1). The interaction term captured this increase directly (change from neutral to valid = .235, 94% highest density interval [.164, .282], posterior probability > 0 = 1), indicating that cueing selectively enhanced evidence accumulation for semantic judgements.

Together, these results suggest a dissociation across decision components, with valid cueing reducing the semantic advantage in non-decision time while simultaneously amplifying the semantic advantage in evidence accumulation, a pattern consistent with cueing differentially affecting preparatory processes prior to decision formation and the efficiency with which information contributes to the decision itself.

Experiment 2

Results from Experiment 1 showed that semantic prioritisation in visual working memory cannot be reduced to a selection bias or just retrieval processes. Although semantic prioritisation is stronger when retrieval is needed, these advantages were attenuated, but not eliminated, when retrieval demands were reduced by a valid retro-cue. DDM further revealed that this pattern reflected dissociable decision components, where semantic information was accessed more rapidly at pre-decisional stages, particularly when retrieval demands were high, while advantages in evidence accumulation were modulated by cueing. Together, these findings indicate that semantic prioritisation reflects a consistent and adaptive reformatting of working memory representations that depends mainly on retrieval demands.

A remaining question is whether this adaptive formatting depends on the need to actively maintain and protect memories over time or from interference⁴⁸. Experiment 2 was designed and pre-registered (<https://researchbox.org/5966>) to address this issue by manipulating both the temporal delay between study and test and the presence of intervening interference, allowing us to examine whether semantic prioritisation is selectively enhanced when visual working memory content must be shielded from competing cognitive demands.

To do so, we designed a task in which we varied time between study and test, and interference of a competing task (see^{49,50} for a similar manipulation). A total of 77 participants took part in the experiment which had three conditions. The *Immediate* condition mirrored the neutral condition of Exp. 1, such that after having studied three items, participants were shown a screen saying “Now”, signalling that retrieval of one of the three items would be commencing immediately. After a short delay (1000ms), the same questions as in Exp. 2 appeared subsequently. In the *Prospective* condition, participants were shown a screen with the word “Later”, whereafter a short delay of 1000ms, they saw a symbol of a triangle, circle or square, together with the word triangle?, circle?, or square? and were asked to respond whether the symbol matched the word (2000ms). After this intervening task, they answered the questions related to one of the items from the study phase. In the *Delayed* condition, participants instead were shown a screen with the word “Later”, whereafter a delay task (1000ms) they were shown a merged symbol of the triangle, circle and square, and were asked to do nothing for the 2000ms it was shown. After this they answered the questions regarding one of the studied items (Fig. 5a).

Critically, the three conditions were designed to dissociate the effects of temporal delay from the need to actively shield working memory representations from interference. Comparing the *Immediate* and *Delayed* conditions isolates the effect of delay in the absence of competing cognitive demands, as no active task intervenes between study and test in either condition. In contrast, comparing the *Delayed* and *Prospective* conditions isolates the effect of interference and active maintenance demands, as both conditions are matched in overall delay but differ in whether participants must protect stored representations while performing a competing task. This design therefore allowed us to assess whether semantic prioritisation is specifically enhanced by the need to shield working memory content from interference, over and above the effects of time alone.

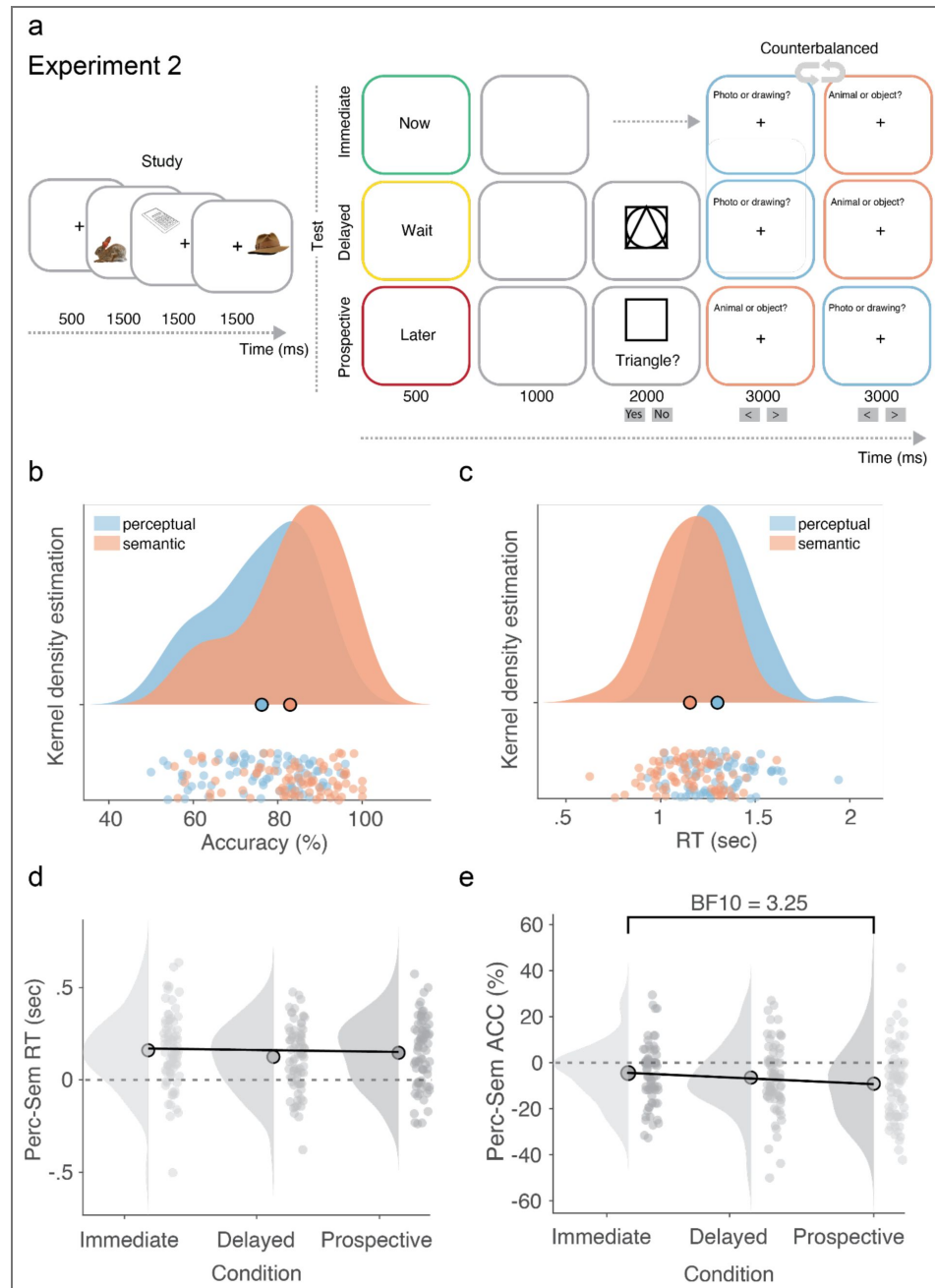


Figure 5. Paradigm and results of Experiment 2.

a The study phase was identical to Exp. 1. After encoding, participants saw three different words ("Now", "Wait", or "Later"), indicating which condition they were in (Immediate, Delayed, or Prospective). In the Immediate condition, they were randomly tested on one of the presented items from the study phase, immediately following a brief 1000ms blank screen. In the Delayed condition, participants were exposed to a symbol consisting of a triangle, square and a circle, stacked on top of each other for 2000ms. After this, they were randomly questioned about one of the studied items. In the Prospective condition, participants saw a triangle, square or a circle with the question "Triangle?", "Square?", or "Circle?", and were asked to respond yes if the word matched the symbol, and no otherwise. After 2000ms, the trial continued in the same manner as the two other conditions. In the Immediate condition, we asked participants the same symbol matching question after they had responded to the questions about the item from encoding (not shown in figure). **b** Average accuracy for semantic trials was higher than perceptual trials. **c** Average reaction times (RTs) for semantic trials were faster than perceptual trials. **d** We observed faster semantic RTs in all conditions, but no difference between conditions. **e** The difference between semantic and perceptual accuracy was lowest for the Immediate condition, followed by Delayed and Prospective, with a significant difference between Immediate and Prospective conditions.

Participants were on average performing better in the semantic question ($M = 82.92\%$, $SD = 11.77\%$) as compared to the perceptual question ($M = 76.22\%$, $SD = 11.71\%$, $t(1,76) = 5.95$, $p < .01$, $BF_{10} > 1000$; Fig 5B). They were also faster at responding to the semantic as compared to the perceptual questions (semantic: $M = 1150\text{ms}$, $SD = 180\text{ms}$; perceptual: $M = 1300\text{ms}$, $SD = 184\text{ms}$; $t(1,76) = 9.49$, $p < .01$, $BF_{10} > 1000$; Fig 5c). Again, semantic features are consistently easier and faster to remember. Moreover, the new experimental conditions introduced in this experiment significantly impacted behaviour (main effect of condition in accuracy: $F = 73$, $p < .001$, $BF > 1000$; RT: $F = 223$, $p < .001$, $BF > 1000$). Participants were most accurate in the Immediate condition ($M = 85.34\%$, $SD = 13.4\%$) compared to the Delayed ($M = 81.08\%$, $SD = 13.66\%$, $BF_{10} > 134$) and the Prospective ones ($M = 72.28\%$, $SD = 14.86\%$, $BF_{10} > 1000$). Interference also impacted behaviour above and beyond the passing of time, as revealed by the greater accuracy in Delayed compared to Prospective ($BF_{10} > 1000$). A similar pattern was observed in RT, with fastest responses in Immediate ($M = 1069\text{ms}$, $SD = 199\text{ms}$) compared to Delayed ($M = 1128\text{ms}$, $SD = 201\text{ms}$, $BF > 1000$) and Prospective ($M = 1396\text{ms}$, $SD = 227\text{ms}$, $BF > 1000$). Prospective yielded slower responses compared to Delayed trials as well ($BF > 1000$).

Our main hypothesis was that if semantic prioritisation is driven by shielding against interfering processes, we would see an increase in semantic prioritisation in the *Prospective* condition as compared to the *Immediate* and *Delayed* condition. Our second hypothesis regarded time between study and test. In our previous paper³⁷, we observed an increase in semantic prioritisation the further away the tested item was from study. However, this measure was confounded by varying the number of items presented at study. Therefore, in the present task, we kept the number of items studied to 3. If time further transformed the memory trace, we would see an increase in semantic prioritisation in the delayed as compared to the immediate condition.

To assess these hypotheses, we conducted a Bayesian repeated-measures ANOVA with Condition (Immediate, Prospective, Delayed) as a within-subject factor, applied separately to reaction time (RT) and accuracy. Bayesian model comparison provided moderate evidence for the null model, indicating no effect of Condition on RT ($BF = .12$). Nevertheless, since we had a pre-registered directional hypothesis, we directly tested *Immediate* vs. *Prospective* and, contrary to our prediction, found strong evidence for the null hypothesis (Bayesian t-test against zero comparing Immediate and Prospective = .09, paired-samples t-test: $t(1,76) = -.51$, $p = .70$, one-sided; difference semantic and perceptual question Immediate condition: $M = 160\text{ms}$, $SD = 198\text{ms}$; Prospective condition: $M = 146\text{ms}$, $SD = 182\text{ms}$; Fig. 5d).

An identical repeated-measures ANOVA, but now using accuracy, showed inconclusive evidence ($BF = .46$). Again, we directly tested the pre-registered directional hypothesis of a difference between *Immediate* and *Prospective* condition. In this case we found moderate evidence for more semantic prioritisation in the prospective as compared to the immediate condition (Bayesian t-test against zero comparing Immediate and Prospective = 3.25; $t(1,76) = 2.35$, $p = .01$, one-sided; difference semantic and perceptual question Immediate condition: $M = -4.5\%$, $SD = 12.61\%$; Prospective condition: $M = -9.13\%$, $SD = 15.77\%$; Fig. 5e), suggesting that a larger delay plus interference from another task increases the semantic prioritisation, reflected in the accuracy of the retrieved memory.

Our second directional hypothesis was that the *Delayed* condition should show more semantic prioritisation as compared to the *Immediate* and *Prospective* condition if time between study and test has an effect on the format of working-memory representations. On the contrary to this prediction, we observed numerically more semantic prioritisation in the *Immediate* ($M = 160\text{ms}$, $SD = 198\text{ms}$) as compared to the *Delayed* ($M = 125\text{ms}$, $SD = 176\text{ms}$) condition when assessing RTs, with the Bayesian t-test against zero showing strong evidence for the null-hypothesis ($BF = .05$; $t(1,76) = -1.46$, $p = .93$, one-sided). When contrasting *Delayed* and *Prospective* conditions, we found strong evidence for the null hypothesis ($BF_{10} = .07$), such that participants had numerically more semantic prioritisation in the *Prospective* condition ($M = 147\text{ms}$, $SD = 182\text{ms}$). For accuracy, we found numerically higher semantic prioritisation in the *Delayed* ($M = -6.46\%$, $SD = 14.50\%$) as compared to the *Immediate* ($M = -4.50\%$, $SD = 12.61\%$). However, Bayesian t-test against zero comparing *Immediate* and *Delayed* showed moderate evidence for the null hypothesis ($BF = .32$,

$t(1,76) = .96$, $p = .17$, one-sided). When contrasting *Delayed* and *Prospective* conditions, we again found numerically more semantic prioritisation in the *Prospective* condition ($M = -9.13\%$, $SD = 15.77\%$), yielding strong evidence for the null hypothesis ($BF_{10} = .06$).

Balanced integration score

We again used BIS to understand the speed-accuracy trade-off. Immediate ($M = 1.57$, $SD = 2.03$) and Delayed ($M = 1.50$, $SD = 2.01$) conditions had numerically less semantic prioritisation than the *Prospective* condition ($M = 1.92$, $SD = 2.12$). Using a one-way Bayesian ANOVA, we found strong evidence for a null effect ($BF_{10} = .015$). Our directional follow-up test testing for a difference in BIS between Immediate and *Prospective* conditions revealed inconclusive evidence ($BF = .50$, $t(1,76) = 1.29$, $p = .1$, one-sided), whereas the test between Immediate and Delayed conditions showed moderate to strong evidence for the null hypothesis ($BF = .10$, $t(1,76) = -.26$, $p = .60$, one-sided).

Across behavioural measures, the results provided limited support for the hypothesis that semantic prioritisation in visual working memory is selectively enhanced by temporal delay alone, but partial support for a role of interference- and delayed-related demands. Reaction time measures showed consistent evidence for the absence of condition effects, with semantic prioritisation in RTs remaining stable across *Immediate*, *Delayed*, and *Prospective* conditions. In contrast, accuracy revealed moderate evidence for increased semantic prioritisation in the *Prospective* condition relative to the *Immediate* condition, suggesting that the presence of an intervening task, requiring the maintenance and protection of memory representations, can enhance the semantic advantage. No reliable differences were observed between *Immediate* and *Delayed* conditions in either RT or accuracy, providing evidence against a simple effect of time between study and test when memory load is held constant. Finally, analyses of the balanced integration score yielded evidence consistent with the null hypothesis, indicating that the observed effects were not driven by systematic speed-accuracy trade-offs. Together, these findings suggest that semantic prioritisation is not uniformly amplified by delay, but may be selectively strengthened when working memory representations must be actively shielded from interference compared to an immediate implementation.

Drift-diffusion modelling of Experiment 2

Together, these findings suggest that interference and time may differentially affect the components of working memory performance, but that these effects are not readily captured by aggregate behavioural measures alone. To characterise the decision processes underlying semantic prioritisation in Experiment 2, we fitted hierarchical drift diffusion models in which feature type (semantic vs perceptual) and condition (Immediate, Prospective, Delayed) jointly modulated non-decision time (t) and drift rate (v). We report results from the joint model allowing both parameters to vary as a function of feature and condition. Again, simpler models along with a more complex model, including threshold as a free parameter are reported in the Supplementary Materials.

Across all conditions, semantic judgements were associated with reliably shorter non-decision times than perceptual judgements, indicating an advantage in processes occurring prior to the onset of evidence accumulation (Fig. 6a). This semantic advantage was present in the *Immediate* condition (mean difference in $t = -.097$ s, 94% highest density interval $[-.117, -.078]$), the *Prospective* condition (mean difference = $-.139$ s, 94% highest density interval $[-.162, -.116]$), and the *Delayed* condition (mean difference = $-.074$ s, 94% highest density interval $[-.097, -.054]$). In all three cases, the posterior probability that the semantic-perceptual difference was negative was equal to or greater than .996, providing strong evidence for a robust semantic access advantage.

Importantly, the magnitude of this advantage differed across conditions. The semantic-perceptual difference in non-decision time was significantly larger in the *Prospective* than in the *Immediate* condition (mean interaction difference = $-.042$ s, 94% highest density interval $[-.073, -.013]$, posterior probability less than zero = .996). The semantic advantage was also reliably larger in the *Prospective* than in the *Delayed* condition (mean difference = $-.065$ s, 94% highest density interval $[-.095, -.037]$, posterior probability less than zero = 1). In contrast, the difference between the

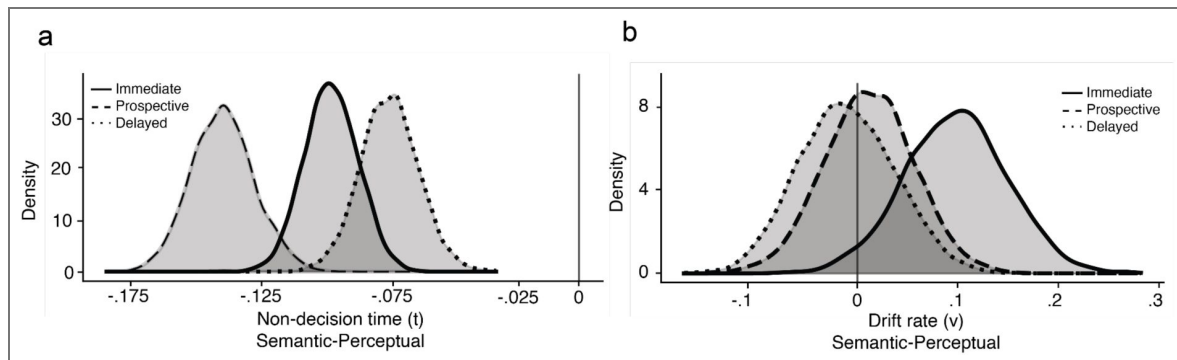


Figure 6. Drift diffusion modelling of Experiment 2.

a We observed a reliable difference in non-decision time between semantic and perceptual questions between Prospective and Immediate and Prospective and Delayed conditions. **b** No such difference was seen in drift rate.

Immediate and *Delayed* conditions was small and uncertain (mean difference = $-.023$ s, 94% highest density interval $[-.052, .004]$). Together, these contrasts indicate that the semantic advantage in non-decision time was selectively enhanced when working memory representations had to be maintained in the face of interference, rather than by increased retention time alone.

Effects of feature type on drift rate were weaker and more condition dependent (Fig. 6b). In the *Immediate* condition, semantic judgements showed a tendency towards higher drift rates than perceptual judgements (mean difference = $.098$), although the 94% highest density interval slightly overlapped zero $[-.000, .191]$. In the *Prospective* condition, the semantic–perceptual difference in drift rate was small and centred near zero (mean difference = $.011$, 94% highest density interval $[-.071, .097]$), and a similar pattern was observed in the *Delayed* condition (mean difference = $-.011$, 94% highest density interval $[-.097, .079]$). A direct comparison between the *Prospective* and *Immediate* conditions revealed a trend towards a reduced semantic drift advantage in the *Prospective* condition (mean interaction difference = $-.086$), but this effect did not reach high posterior certainty (94% highest density interval $[-.212, .046]$). No reliable differences in semantic drift advantages were observed between the *Prospective* and *Delayed* or between the *Immediate* and *Delayed* conditions.

Together, and in line with Experiment 1, the drift diffusion modelling results indicate that semantic prioritisation in Experiment 2 primarily reflects differences in processes occurring prior to evidence accumulation rather than differences in evidence accumulation itself. Semantic judgements were associated with shorter non-decision times than perceptual judgements across all conditions, with this advantage amplified when memories had to be protected from interference in the *Prospective* condition. Thus, prospective cues altered preparatory processing prior to evidence accumulation, rather than the accumulation process itself.

Discussion

Across experiments, semantic judgments showed a reliable advantage in non-decisional processing time. This indicates that semantic information confers an advantage at stages preceding evidence accumulation, while differences in evidence accumulation itself depended on attentional context. These findings are consistent with, but do not uniquely demonstrate, more efficient access to semantic representations in working memory and supports the view that working memory representations are dynamically transformed as attentional demands increase^{5,6}. Crucially, this semantic prioritisation was expressed primarily as a reduction in non-decision time, rather than as a uniform speeding of responses. Moreover, the magnitude of the effect was modulated by task demands and could not be explained by a fixed response bias. Instead, the semantic access advantage was selectively amplified when items fell outside the focus of attention, particularly under conditions of attentional disruption.

Reanalysis of our earlier dataset using drift diffusion modelling revealed a robust reduction in non-decision time for semantic judgements across memory loads and lags, consistent with semantic information influencing decision formation earlier following probe onset. By contrast, effects on drift rate were contingent on task demands, with semantic accumulation advantages emerging most clearly under higher load and longer lags. This dissociation suggests a functional separation between processes governing the timing of information availability and those governing the efficiency of evidence accumulation. Notably, this pattern mirrors distinctions drawn in the episodic memory literature, where semantic or conceptual information often precedes perceptual detail during retrieval, consistent with top-down or reconstructive access dynamics rather than feedforward perceptual processing^{30,31}. However, on the basis of this dataset alone, reduced non-decision time cannot disambiguate whether semantic information becomes available earlier due to differences in retrieval dynamics or later-stage decision-related processes, leaving open which component of the decision process primarily drives the observed semantic advantage.

Experiment 1 was designed to test whether semantic prioritisation is differentially expressed when selection uncertainty at test is reduced: whether semantic information was more readily usable for guiding behaviour following probe onset, or whether it is merely weighted more

strongly at the decision stage when attention can be strategically deployed. Behaviourally, semantic prioritisation increased when participants could not pre-select a single item and instead had to maintain multiple representations concurrently, retrieving the relevant one only at probe. Drift-diffusion modelling revealed that under neutral cueing, non-decision time was selectively prolonged for perceptual judgements, whereas the semantic advantage in non-decision time remained relatively unaffected, indicating a robust semantic advantage at pre-decisional stages when attentional guidance was unavailable. By contrast, cueing prior to the probe selectively modulated drift rate, showing that attentional prioritisation also alters how efficiently retrieved information is translated into a decision. Whereas previous work has primarily attributed retro-cueing effects to changes in representational quality or precision, often indexed by drift-rate modulations¹⁴, the present findings indicate that semantic prioritisation is expressed primarily as an access advantage. This dissociation suggests that internal attention and retrieval demands act on distinct stages of working memory decision processes. Together, these results show that semantic prioritisation is not a fixed decision bias, but an adaptive property of working memory that becomes most pronounced when attentional guidance is uncertain, shaping both how information is accessed and how it is subsequently used to guide behaviour.

In contrast to our previous findings³⁷, we did not observe a relationship between semantic prioritisation and lag in the neutral condition. In the earlier study, lag was necessarily confounded with memory load, such that increasing temporal distance also implied that individual items spent more time outside the focus of attention. Although Experiment 1 revealed lag-related increases in semantic prioritisation in the valid condition, most clearly in accuracy and the balanced integration score, this effect was absent when retrieval demands were high and attention had to remain distributed across items. When this confound was explicitly addressed by holding memory load constant while independently manipulating delay and interference, we found no evidence that temporal delay alone modulated semantic prioritisation.

One possible explanation for this null effect comes from work by Oberauer¹⁸, who showed that relatively short retention intervals of around two seconds are sufficient to separate storage from concurrent processing in working memory. Under such conditions, information can be maintained in a passive state without remaining in the focus of attention. In our Delay condition, the maintenance interval may therefore have been long enough for participants to disengage attentional processing from the stored items, effectively decoupling retention from ongoing cognitive operations. As a result, temporal delay alone may not have increased the likelihood that representations fell outside the focus of attention in a way that would further amplify semantic prioritisation. This interpretation aligns with models in which representations outside the focus of attention can be maintained without ongoing processing demands, such that delay alone does not necessarily increase retrieval difficulty unless attentional resources are actively diverted^{18,51}.

Experiment 2 further dissociated the influence of temporal delay from that of attentional disruption during the maintenance interval. Semantic prioritisation was selectively enhanced when working memory representations had to be maintained in the face of attentional disruption, whereas delay in the absence of interference produced weaker and less reliable effects. Crucially, this dissociation was expressed most robustly in non-decision time, indicating that attentional disruption alters how rapidly semantic versus perceptual information can be accessed prior to evidence accumulation. This pattern suggests that the critical driver of semantic prioritisation is not temporal decay per se, but the need to preserve and later reinstate task-relevant information under conditions of limited attentional support. Consistent with this interpretation, a substantial body of work indicates that interference, rather than the mere passage of time, is a primary determinant of forgetting and representational degradation in working memory^{44,45,48}. Together, these findings suggest that semantic prioritisation reflects an adaptive response to attentional disruption rather than a passive consequence of retention interval. In addition, these results also indicate that retro-cue benefits in VWM cannot be understood in terms of stronger representational precision or maintenance. Instead, internal attention seems to impact the stage at which information becomes available for decision-making.

Taken together, the results support an account in which semantic representations are more robust than perceptual representations to attentional misrouting and interference. When attention cannot be dedicated to a single item, working memory performance increasingly depends on representational formats that can be rapidly re-accessed and flexibly deployed for decision making. Semantic information appears to be privileged under these conditions. A natural implication is that, as items fall outside the focus of attention, their functional format may shift towards more abstract, memory-like representations that are resilient to disruption, at the expense of fine-grained perceptual detail ^{23,52}.

This interpretation is consistent with influential models that distinguish a narrow focus of attention from a broader set of activated but unattended representations in working memory ^{17,18,51}, as well as with accounts proposing that unattended or activity-silent working memory relies on mechanisms that overlap computationally with episodic memory retrieval ^{21,23,24}. Within this framework, semantic prioritisation reflects the relative resilience of conceptual information when representations must be accessed from an unattended state, rather than a default property of working memory per se.

It is important to highlight that because non-decision time aggregates encoding of the cue, access, and motor components, the present results cannot uniquely attribute the semantic advantage to purely retrieval latency. Instead, the findings constrain the stage of processing at which semantic advantages arise. Future work manipulating encoding and response preparation independently will be required to further isolate the underlying mechanism.

More broadly, our findings suggest that the limits of working memory are not defined solely by storage capacity, but by which representational formats remain rapidly accessible when attentional support is uncertain. Semantic information is favoured in this regime because it remains usable for guiding behaviour under conditions of attentional disruption. Importantly, this framework links working memory, episodic memory, and imagery by highlighting a shared reliance on retrieval-based access mechanisms when representations are internally generated or no longer sustained by attention ^{21,24,35,36}. From this perspective, semantic prioritisation reflects not a specialised feature of working memory alone, but a more general principle governing how information is accessed across memory systems.

Method

Reanalysis of Kerrén et al. 2022

We reanalysed data from ³⁷, in which participants ($n = 103$ after exclusions) performed a visual working memory task requiring maintenance of one to four object images across a short delay. Each object varied orthogonally in semantic category (animate/inanimate) and visuo-perceptual format (photograph/drawing). At the test, participants answered either a semantic or a perceptual question about one studied item, where reaction time and accuracy were obtained. Memory load and the lag between encoding and test were systematically varied. The present analyses focus on Experiment 2, which included both load and lag manipulations. Full procedural details are reported in the original study. Here, we applied hierarchical drift diffusion modelling to dissociate pre-decisional access from evidence accumulation processes.

Experiment 1

Participants

We recruited 96 participants online (48 females, 47 males, and 1 did not prefer to say) with a mean age of 29.69 years ($SD \pm 4.18$ years, range = 21-35 years). Participation was with informed consent via the Prolific platform (<https://www.prolific.ac/> ³⁸). The eligibility criteria were that participants had to be between 18 and 35 years old, fluent in English, being from the United Kingdom or the United States, normal, or corrected-to-normal vision, have a minimum of 30 previous survey submissions, and have a minimum approval rate of 95 on Prolific. Participants were reimbursed with £4.5 for taking part in the experiment. The experiment lasted approximately 30 minutes. The experiment was terminated prematurely for participants who failed to pass attention checks (see

below), which resulted in no payment ($n = 0$ participants). We also excluded participants whose accuracy did not significantly exceed chance level ($p < .05$, Binomial test against 50% correct responses) in either the semantic or perceptual judgments ($n = 13$ participants). Thus, 83 participants remained for analysis. The experiment was approved by the University of Granada Ethics Committee. Sample size and exclusion criteria was pre-registered before data collection (see <https://researchbox.org/5966> [↗](#)).

Stimuli

The experimental stimuli comprised 150 pictures (75 depicting animals, 75 depicting inanimate objects). An additional 10 pictures were used for instructions and practice trials. A detailed description of the stimulus set is provided in [31](#). For each of the pictures (which are color photographs), a line-drawing version (black-and-white) was created using GNU image manipulation software (<http://www.gimp.org> [↗](#)). For each participant, half of the animals and half of the inanimate objects were shown as line drawings (randomly assigned) and the others as color photographs. The stimuli thus differed in two orthogonal dimensions, visuo-perceptual (photograph/drawing) and semantic (animate/inanimate; [Fig. 1a](#) [↗](#)).

Task

Each trial started with a central fixation cross (500ms), after which 3 pictures were sequentially presented (1500ms each) at different random locations on the screen ([Fig. 1b](#) [↗](#)). The locations were randomly selected from 16 equidistant positions on an invisible circle around screen centre. The positions were selected to be at minimum 45° apart from each other. Participants were asked to remember the pictures and locations. Pictures were pseudo-randomly chosen so that they never share the same were all of the same (perceptual or semantic features.) type. After a 1000ms blank screen, a retro-cue was shown for 500ms. In the *valid cue* condition, a circle was shown in the location where the questions would later appear. In the *neutral cue* condition, three circles were shown in the three locations that items were previously studied. After a delay (1000ms), a question appeared at the location of one of the objects. The question was aimed at either probing the semantic (“Animal or Object?”) or perceptual (“Photo or Drawing?”) features of the studied item that was presented at that location. These questions were counterbalanced within participants such that the semantic and perceptual questions were equally likely to be asked first. In addition, across participants we also counterbalanced the order of words in the question, (e.g., “Object or Animal?” and “Drawing or Photo?”) and how these different options were tested together (e.g., if “Object or Animal?” and “Drawing or Photo?” were the two response options for one participant, then “Animal or Object?” and “Drawing or Photo?” could be the response options for the next participant and so on). Participants were asked to select the response option to report its characteristics (perceptual and semantic) using the right and left arrow keys. Participants were given 3 seconds to make their selections (where they could not click a button during the first 200ms). Upon selection the font colour of the question turned red (for incorrect) or green (for correct). Independent if participants responded or not within 3 seconds of the first question, the next question appeared (if the first one probed semantic features, the second probed perceptual features and vice versa). Again, independent of engagement from participants, a new trial began after 3 seconds. All trials with no response were excluded from analysis.

Procedure

After extensive written instructions, participants performed three practice trials to familiarise themselves with the task. Participants were free to repeat the instructions and practice trials until they felt confident at performing the task. Thereafter, each participant performed 100 experimental trials (50 in each condition). After the 4th, 16th, and 28th trial, a brief attention check was performed. For this, a large filled triangle was presented centrally in either green, blue, or red color and participants were asked to name the colour via button press. If a participant failed on any of the attention checks, the experiment was aborted (see above, *Participants*). After 50 trials, participants could take a short break (self-paced).

Experiment 2

Participants

We recruited 110 participants online (64 female, 45 male, and 1 who did not prefer to say) with a mean age of 29.87 years ($SD \pm 4.28$ years, range = 20-35 years). The task took approximately 35 minutes and participants were compensated £6 for taking part in the experiment. Again, we used attention checks, and the experiment was terminated prematurely for participants who failed to pass attention checks, which resulted in no payment ($n = 1$ participants). We also excluded participants whose accuracy did not significantly exceed chance level ($p < .05$, Binomial test against 50% correct responses) in either the semantic or perceptual judgments ($n = 33$ participants). Thus, 77 participants remained for analysis. Sample size and exclusion criteria was pre-registered before data collection (see <https://researchbox.org/5966> ) . The experiment was approved by the University of Granada Ethics Committee.

Stimuli, Task, and Procedure

The design of Exp. 2 was similar to Exp. 1, with the following differences. After having studied 3 items, at test, a cue signalling which condition they were in, was shown for 500ms with a delay of 1000ms both preceding and following the condition-cue. The cue was “Now”, “Later” or “Wait”, depending on the condition they were in. If the cue was “Now” (*Immediate* condition) it meant that the upcoming questions would be administered immediately following the delay (equal to the *neutral* condition in Exp. 1). If the cue was “Later” (*Prospective* condition), participants were first exposed to a symbol (triangle, square or circle, randomised across trials) and a question (triangle?, square?, circle?, randomised across trials), and the response options (yes / no, or no / yes, counterbalanced). They had 2000ms to respond, whereafter the questions regarding the studied items appeared. If the cue was “Wait” (*Delayed* condition), a symbol made up of stacking the triangle, circle and square on top of each other was shown for 2000ms. The task for the participants was to just observe the symbol, whereafter the questions regarding the studied items appeared. To equate cognitive load in the three conditions, in the immediate condition, the questions regarding the symbols appeared after participants had responded to the questions regarding the studied items. Participants performed a total of 100 trials, divided into 34, 33, and 33 trials in each condition.

Data analysis

The BIS was calculated by standardising (z-scoring) the RTs on correct trials and the accuracies relative to their mean and standard deviation over all conditions, and then subtracting within each condition the RT score from the accuracy score ⁴⁷. Main analyses were performed in MATLAB, version 2022b (The MathWorks, Munich, Germany) and Bayesian ANOVAs were performed on JASP 0.95.4 using default priors and computing effects across matched models.

Drift diffusion modelling

We fitted hierarchical drift diffusion models (HDDMs) to single-trial choice and reaction time data using the *HSSM package* in Python ⁵³. Analyses were restricted to reaction times greater than 0.2 s. Outliers were addressed using a mixture modeling approach (via the `p_outlier` parameter in *HSSM*), which assumes data arise from a mixture of the intended process and a uniform “noise” distribution with a fixed probability of 0.05, meaning that 5% of trials were treated as outliers. The critical predictor was Feature, indicating whether the probe required a perceptual judgement (photo versus drawing) or a semantic judgement (animate versus object). All models used the standard DDM parameterisation and were estimated with Bayesian inference. We used hierarchical (partial pooling) random intercepts by participants for the modelled parameters, allowing parameters to vary across individuals while sharing information at the group level. For each model, we ran four independent chains using the *NUTS* (NumPyro backend) sampler, each consisting of 2000 total iterations; the first 1000 samples were discarded as burn-in, to allow the algorithm to converge toward the posterior distribution, leaving 1000 draws per chain for a total

of 4000 retained samples. Comparison between models was done using the *Arviz* package in Python (*arviz.loo*), computing the Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO-CV).

Model 1

Model 1 tested whether semantic judgements differed from perceptual judgements in non-decision time (t), which captures processes outside evidence accumulation, including stimulus encoding and motor execution. In this model, t was specified as a function of Feature with participant-level random intercepts:

$$t \sim \text{Feature} + (1 | \text{participant})$$

All other DDM parameters were held constant across features at the group level. The key quantity of interest was the posterior for the Feature coefficient on t , reflecting a change in non-decision time for semantic relative to perceptual judgements.

Model 2

Model 2 tested whether Feature modulated the drift rate (v), which reflects the efficiency of evidence accumulation towards the correct response. Drift rate was specified as a function of Feature with participant-level random intercepts:

$$v \sim \text{Feature} + (1 | \text{participant})$$

All other parameters were held constant across features at the group level. The key quantity of interest was the Feature effect on v , reflecting faster or slower accumulation for semantic relative to perceptual judgements.

Model 3

Model 3 allowed Feature to affect both drift rate and non-decision time within a single model, testing whether semantic advantages reflect changes in pre-decisional processes, evidence accumulation, or a combination of both. Both parameters were modelled as functions of Feature with participant-level random intercepts:

$$v \sim \text{Feature} + (1 | \text{participant})$$

$$t \sim \text{Feature} + (1 | \text{participant})$$

This joint model assesses the extent to which any semantic advantage is attributable to faster access and encoding processes (lower t), more efficient evidence accumulation (higher v), or both.

Model 4

Model 4 extended the joint model by additionally allowing Feature to modulate the decision threshold, testing whether semantic prioritisation also reflects differences in response caution or decision policy. Drift rate, non-decision time, and threshold were all modelled as functions of Feature, with participant-level random intercepts:

$$v \sim \text{Feature} + (1 | \text{participant})$$

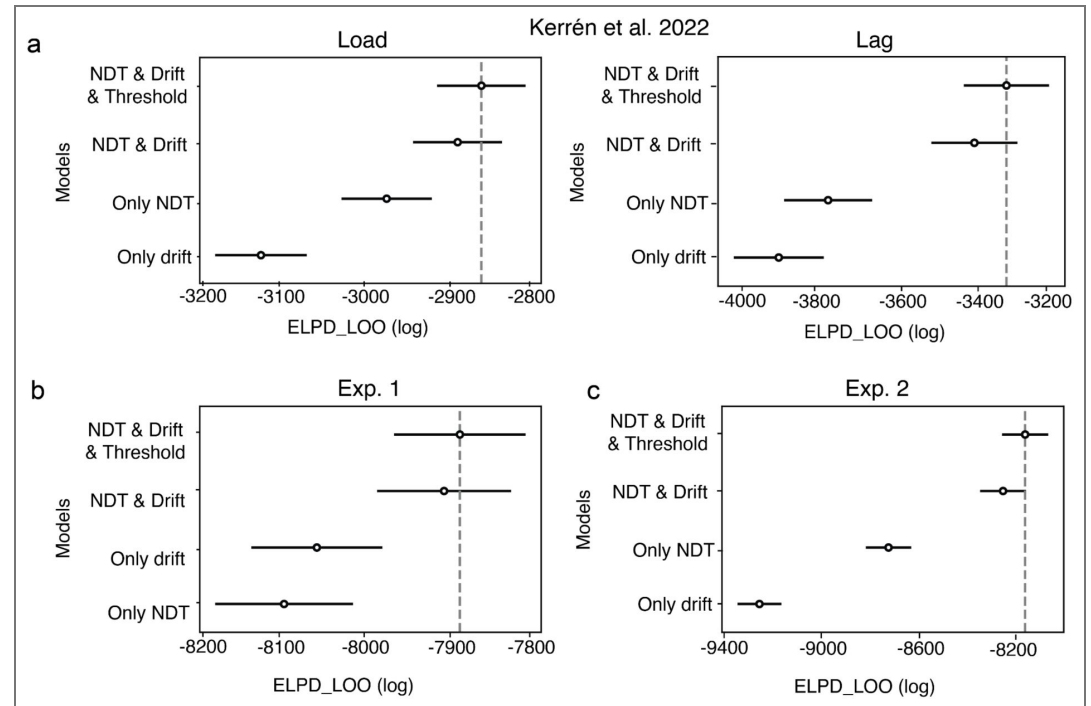
$$t \sim \text{Feature} + (1 | \text{participant})$$

$$a \sim \text{Feature} + (1 | \text{participant})$$

This model therefore assesses whether any semantic advantage can be attributed not only to differences in the timing with which information becomes available for decision-making (non-decision time) or the efficiency of evidence accumulation (drift rate), but also to differences in response caution, as indexed by the amount of evidence required to commit to a response (threshold).

In all models, response bias was set to $z = .5$, assuming no bias for a given boundary in semantic or perceptual questions.

Supplementary material



Supplementary Figure 1. Comparison of the 4 drift-diffusion models used. **a** The most complex model showed the best fit to the data followed by a model with non-decision time and drift as free parameters when analysing load 3 (left) and lag 3 (right). **b** The same was true in Experiment 1. **c** and Experiment 2.

Data availability

The code to reproduce the results of this study is available at https://github.com/kerrencasper/Flipper_2025_git

Acknowledgements

C.K. is supported by the Max Planck society. C.G.G. is supported by Project PID2023-149428NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by FEDER, EU, and Grant RYC2021-033536-I funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGeneration EU/PRTR. J.L.D. is supported by Project PID2023-151104NA-I00 funded by MCIN/AEI/10.13039/501100011033 and by FEDER, EU, and Grant RYC2021-033940-I funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGeneration EU/PRTR. The Mind, Brain and Behavior Research Center receives funding from grants CEX2023-001312-M by MICIU/AEI/10.13039/501100011033 and UCE-PP2023-11 by the University of Granada. This work was supported by the Unidad de Excelencia María de Maeztu – CIMCYC (CEX2023-001312-M), funded by MICIU/AEI/10.13039/501100011033. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Additional information

Author contributions

C.K., J.L.D. and C.G.G. were responsible for conceptualisation, project administration, resources, data collection, data curation, and formal analysis. All authors were involved in methodology, visualisation, and writing the original draft, reviewing, and editing.

Funding

Funder	Grant reference number	Author
Ministerio de Ciencia e Innovación (MCIN)	PID2023-149428NB-I00	Carlos González-García
		Juan Linde-Domingo
Ministerio de Ciencia e Innovación (MCIN)	RYC2021-033536-I	Juan Linde-Domingo
		Carlos González-García

Author ORCID iDs

Casper Kerrén:  <https://orcid.org/0000-0003-4870-6072>

References

1. **Baddeley A** (1992) Working memory. *Science* **255**:556-559 <https://doi.org/10.1126/science.1736359> | PubMed
2. **Baddeley A** (2012) Working memory: theories, models, and controversies. *Annu Rev Psychol* **63**:1-29 <https://doi.org/10.1146/annurev-psych-120710-100422> | PubMed
3. **D'Esposito M, Postle BR** (2015) The cognitive neuroscience of working memory. *Annu Rev Psychol* **66**:115-142 <https://doi.org/10.1146/annurev-psych-010814-015031> | PubMed
4. **Pasternak T, Greenlee MW** (2005) Working memory in primate sensory systems. *Nat Rev Neurosci* **6**:97-107 <https://doi.org/10.1038/nrn1603> | PubMed
5. **Harrison SA, Tong F** (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458**:632-635 <https://doi.org/10.1038/nature07832> | PubMed

6. Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci* **20**:207-214 <https://doi.org/10.1111/j.1467-9280.2009.02276.x> | PubMed
7. Spaak E, Watanabe K, Funahashi S, Stokes MG (2017) Stable and Dynamic Coding for Working Memory in Primate Prefrontal Cortex. *J Neurosci* **37**:6503-6516 <https://doi.org/10.1523/jneurosci.3364-16.2017> | PubMed
8. Rademaker RL, Chunharas C, Serences JT (2019) Coexisting representations of sensory and mnemonic information in human visual cortex. *Nat Neurosci* **22**:1336-1344 <https://doi.org/10.1038/s41593-019-0428-x> | PubMed
9. Albers AM, Kok P, Toni I, Dijkerman HC, de Lange FP (2013) Shared representations for working memory and mental imagery in early visual cortex. *Curr Biol* **23**:1427-1431 <https://doi.org/10.1016/j.cub.2013.05.065> | PubMed
10. Kwak Y, Curtis CE (2022) Unveiling the abstract format of mnemonic representations. *Neuron* **110**:1822-1828.e5 <https://doi.org/10.1016/j.neuron.2022.03.016> | PubMed
11. Pettini L, et al. (2025) Visual working memory representations of naturalistic images in early visual cortex are not sensory-like. *bioRxiv* <https://doi.org/10.64898/2025.12.11.693774>
12. Myers NE, Stokes MG, Nobre AC (2017) Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends Cogn Sci* **21**:449-461 <https://doi.org/10.1016/j.tics.2017.03.010> | PubMed
13. Linde-Domingo J, Spitzer B (2024) Geometry of visuospatial working memory information in miniature gaze patterns. *Nat Hum Behav* **8**:336-348 <https://doi.org/10.1038/s41562-023-01737-z> | PubMed
14. Cipriani GA, González-García C, Martín-Arévalo E, Lupiáñez J, Botta F (2024) Competing working memory contents: Perceptual over semantic prioritization and voluntary retrieval following retro-cueing. *Visual Cognition* **32**:342-361 <https://doi.org/10.1080/13506285.2024.2435652>
15. Jensen O, Lisman JE (1996) Novel lists of 7 +/-2 known items can be reliably stored in an oscillatory short-term memory network: interaction with long-term memory. *Learn. Mem* **3**:257-263 <https://doi.org/10.1101/lm.3.2-3.257> | PubMed
16. Miller GA (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* **63**:81-97 <https://doi.org/10.1037/0033-295x.101.2.343> | PubMed
17. Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci* **24**:87-114 <https://doi.org/10.1017/s0140525x01003922> | PubMed
18. Oberauer K (2002) Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **28**:411-421 <https://doi.org/10.1037/0278-7393.28.3.411>
19. Olivers CNL, Peters J, Houtkamp R, Roelfsema PR (2011) Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn Sci* **15**:327-334 <https://doi.org/10.1016/j.tics.2011.05.004> | PubMed
20. Mongillo G, Barak O, Tsodyks M (2008) Synaptic theory of working memory. *Science* **319**:1543-1546 <https://doi.org/10.1126/science.1150769> | PubMed
21. Stokes MG (2015) 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Sciences* **19**:394-405 <https://doi.org/10.1016/j.tics.2015.05.004> | PubMed
22. Jeneson A, Squire LR (2012) Working memory, long-term memory, and medial temporal lobe function. *Learn. Mem* **19**:15-25 <https://doi.org/10.1101/lm.024018.111> | PubMed
23. Beukers AO, Buschman TJ, Cohen JD, Norman KA (2021) Is Activity Silent Working Memory Simply Episodic Memory?. *Trends in Cognitive Sciences* **25**:284-293 <https://doi.org/10.1016/j.tics.2021.01.003> | PubMed

24. Rose NS, et al. (2016) Reactivation of latent working memories with transcranial magnetic stimulation. *Science* **354**:1136-1139 <https://doi.org/10.1126/science.aah7011> | PubMed
25. DiCarlo JJ, Zoccolan D, Rust NC (2012) How Does the Brain Solve Visual Object Recognition?. *Neuron* **73**:415-434 <https://doi.org/10.1016/j.neuron.2012.01.010> | PubMed
26. Serre T, Oliva A, Poggio T (2007) A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. U.S.A* **104**:6424-6429 <https://doi.org/10.1073/pnas.0700622104> | PubMed
27. Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time. *Nat Neurosci* **17**:455-462 <https://doi.org/10.1038/nn.3635> | PubMed
28. Carlson T, Tovar DA, Alink A, Kriegeskorte N (2013) Representational dynamics of object vision: the first 1000 ms. *J Vis* **13** <https://doi.org/10.1167/13.10.1> | PubMed
29. Clarke A, Tyler LK (2015) Understanding What We See: How We Derive Meaning From Vision. *Trends in Cognitive Sciences* **19**:677-687 <https://doi.org/10.1016/j.tics.2015.08.008> | PubMed
30. Lifanov J, Linde-Domingo J, Wimber M (2021) Feature-specific reaction times reveal a semanticisation of memories over time and with repeated remembering. *Nat Commun* **12**:3177 <https://doi.org/10.1038/s41467-021-23288-5> | PubMed
31. Linde-Domingo J, Treder MS, Kerrén C, Wimber M (2019) Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat Commun* **10** <https://doi.org/10.1038/s41467-018-08080-2> | PubMed
32. Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review* **110**:611-646 <https://doi.org/10.1037/0033-295x.110.4.611> | PubMed
33. Kerrén C, Zhao Y, Griffiths BJ (2024) A reduction in self-reported confidence accompanies the recall of memories distorted by prototypes. *Commun Psychol* **2**:1-11 <https://doi.org/10.1038/s44271-024-00108-2> | PubMed
34. Kerrén C, Reznik D, Doeller CF, Griffiths BJ (2025) Exploring the role of dimensionality transformation in episodic memory. *Trends in Cognitive Sciences* <https://doi.org/10.1016/j.tics.2025.01.007> | PubMed
35. Dijkstra N, Mostert P, Lange FP, Bosch S, van Gerven MA (2018) Differential temporal dynamics during visual imagery and perception. *eLife* **7**:e33904 <https://doi.org/10.7554/eLife.33904> | PubMed
36. Dijkstra N, Ambrogioni L, Vidaurre D, van Gerven M (2020) Neural dynamics of perceptual inference and its reversal during imagery. *eLife* **9**:e53588 <https://doi.org/10.7554/eLife.53588> | PubMed
37. Kerrén C, Linde-Domingo J, Spitzer B (2022) Prioritization of semantic over visuo-perceptual aspects in multi-item working memory. *bioRxiv* <https://doi.org/10.1101/2022.06.29.498168>
38. Phylactou P, Traikapi A, Papadatou-Pastou M, Konstantinou N (2022) Sensory recruitment in visual short-term memory: A systematic review and meta-analysis of sensory visual cortex interference using transcranial magnetic stimulation. *Psychon Bull Rev* **29**:1594-1624 <https://doi.org/10.3758/s13423-022-02107-y> | PubMed
39. van Ede F, Nobre AC (2023) Turning Attention Inside Out: How Working Memory Serves Behavior. *Annu Rev Psychol* **74**:137-165 <https://doi.org/10.1146/annurev-psych-021422-041757> | PubMed
40. Griffin IC, Nobre AC (2003) Orienting attention to locations in internal representations. *J Cogn Neurosci* **15**:1176-1194 <https://doi.org/10.1162/089892903322598139> | PubMed
41. Ratcliff R (1978) A theory of memory retrieval. *Psychological Review* **85**:59-108 <https://doi.org/10.1037/0033-295x.85.2.59>
42. Ratcliff R, McKoon G (2008) The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput* **20**:873-922 <https://doi.org/10.1162/neco.2008.12-06-420> | PubMed
43. Wiecki TV, Sofer I, Frank MJ (2013) HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front Neuroinform* **7** <https://doi.org/10.3389/fninf.2013.00014> | PubMed

44. **Portrat S**, Barrouillet P, Camos V (2008) Time-related decay or interference-based forgetting in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition* **34**:1561-4 <https://doi.org/10.1037/a0013356> | [PubMed](#)
45. **Ricker TJ**, Cowan N, Morey CC (2010) Visual working memory is disrupted by covert verbal retrieval. *Psychon Bull Rev* **17**:516-521 <https://doi.org/10.3758/pbr.17.4.516> | [PubMed](#)
46. **Kuo B.-C**, Stokes MG, Nobre AC (2012) Attention Modulates Maintenance of Representations in Visual Short-term Memory. *J Cogn Neurosci* **24**:51-60 https://doi.org/10.1162/jocn_a_00087 | [PubMed](#)
47. **Liesefeld HR**, Fu X, Zimmer HD (2015) Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *J Exp Psychol Learn Mem Cogn* **41**:1140-1151 <https://doi.org/10.1037/xlm0000081> | [PubMed](#)
48. **Oberauer K**, Lin H.-Y (2017) An interference model of visual working memory. *Psychol Rev* **124**:21-59 <https://doi.org/10.1037/rev0000044> | [PubMed](#)
49. **Padezhki I**, Linde-Domingo J, González-García C (2025) Dynamic prioritization reshapes neural geometries for action in human working memory. *bioRxiv* <https://doi.org/10.1101/2025.07.18.665374>
50. **van Loon AM**, Olmos-Solis K, Fahrenfort JJ, Olivers CN (2018) Current and future goals are represented in opposite patterns in object-selective cortex. *eLife* **7**:e38677 <https://doi.org/10.7554/eLife.38677> | [PubMed](#)
51. **Oberauer K**, Hein L (2012) Attention to information in working memory. *Current Directions in Psychological Science* **21**:164-169 <https://doi.org/10.1177/0963721412444727>
52. **Zhang W**, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* **453**:233-235 <https://doi.org/10.1038/nature06860> | [PubMed](#)
53. **Fengler A**, Xu Y, Bera K, Omar A, Frank MJ, Fengler A, Xu P, Bera K, Omar A, Frank MJ (2025) HSSM: A generalized toolbox for hierarchical Bayesian estimation of computational models in cognitive neuroscience (in preparation).

Peer reviews

Reviewer #1 (Public review):

Summary:

This paper investigates whether semantic prioritization in visual working memory reflects pre-decisional access, evidence accumulation, or both, using drift diffusion modeling across a reanalysis of prior data and two new experiments. The core finding - that semantic information receives a robust pre-decisional access advantage that is amplified by attentional disruption rather than temporal delay alone - is novel and contributes meaningfully to ongoing debates about the format and accessibility of working memory representations.

Strengths:

The experimental approach is well-motivated, and the use of drift-diffusion modeling to decompose decision components adds analytical value beyond standard RT and accuracy measures. The two new experiments are pre-registered and address important questions. The broader theoretical conclusion - that working memory limits are shaped not only by storage capacity but by which representational formats remain accessible under attentional uncertainty - is an important and timely contribution to the field.

Weaknesses:

The central interpretive claims rely heavily on differences in non-decision time, a parameter that aggregates many processes unrelated to memory retrieval, making it rather difficult to uniquely attribute the observed effects to access or retrieval mechanisms specifically.

Additionally, the characterization of the two memory conditions as genuinely perceptual versus semantic warrants further justification, as both may primarily require categorical rather than format-specific knowledge.

<https://doi.org/10.7554/eLife.111859.1.sa2>

Reviewer #2 (Public review):

This manuscript aims to characterize how semantic information is prioritized relative to perceptual details in visual working memory. The central claim is that semantic judgements benefit from faster pre-decisional access (shorter non-decision time), and that advantages in evidence accumulation emerge under higher cognitive demands (e.g., when items are outside the focus of attention or must be maintained under interference). Based on this, the paper argues that unattended working-memory contents are reformatted into more abstract, long-term-memory-like semantic representations that remain more readily accessible than fine-grained perceptual features.

Strengths:

- (1) The question is timely and relevant to current research about the format of visual working memory.
- (2) Behaviorally, the semantic advantage is carefully documented in many conditions across datasets.
- (3) The use of hierarchical drift-diffusion modelling is helpful to decompose the semantic advantage into cognitive processes such as non-decision time and drift-rate components.

Weaknesses:

- (1) The strong claims about visual working-memory representation and "long-term-memory-like" formats rest on an indirect inference from decision-model parameters to representational content, and this link is not convincingly established. Non-decision time, as implemented here, bundles many things, such as probe processing, cue processing, retrieval/access, and motor preparation, so reduced non-decision time for semantic probes could reflect easier question reading, simpler response mapping, or more efficient decision preparation rather than a genuine advantage in accessing semantic memory representations. Although the manuscript acknowledges that non-decision time includes multiple processes, it nonetheless treats this parameter as primary evidence for a retrieval-stage semantic advantage, which overstates what the data can uniquely support.
- (2) The modelling approach is relatively constrained and does not fully address the underdetermination inherent in mapping latent drift-diffusion parameters onto specific psychological mechanisms. The preferred model that allows multiple parameters (non-decision time, drift rate, threshold) to vary provides only modest improvements in predictive accuracy over simpler models, and several key drift-rate effects are present only in particular load or lag conditions. As a result, the theoretical interpretation that semantic prioritization primarily reflects faster access and secondarily more efficient accumulation under high demand appears rather post hoc, and alternative accounts focused on generic task efficiency or strategy differences remain plausible.
- (3) The operationalization of "semantic" is narrow and largely categorical, focusing on animacy (animal/object) and a perceptual format dimension (photo/drawing), rather than richer semantic or associative relations among items. This makes it difficult to generalize the conclusions to broader claims about semantic structure and its integration into working-memory representations. Important recent work on how semantic and associative relationships facilitate the formation, maintenance, and retrieval of visual working memory

is not adequately integrated into the theoretical framing. Consequently, the discussion tends to generalize from a specific probe structure to a broader semantic prioritization theory without engaging fully with the existing literature on semantic facilitation and neural decoding of working-memory content.

(4) The paper contrasts its behavioral/model-based results with prior neural decoding findings, but the comparison is not fair. Neural decoding provides complementary evidence about the content and format of working memory representations, whereas drift-diffusion parameters reflect downstream decision dynamics given a probe. Because the current work does not include any direct representational or neural measure, its conclusions about representational "reformatting" and long-term-memory-like access remain speculative and, in places, feel like a stretch.

(5) Overall, while the data show a semantic advantage in decision-stage measures and the modelling provides an informative decomposition of this advantage, the manuscript does not fully achieve its stated aim of characterizing the representational format of visual working memory or demonstrating a mechanistic shift toward long-term-memory-like semantic representations. The work primarily informs decision-process analyses of the conditions under which semantic judgements are faster and more robust, rather than the nature of visual working-memory representations themselves.

<https://doi.org/10.7554/eLife.111859.1.sa1>

Author response:

We thank the reviewers for their constructive and careful assessment of our manuscript. We are encouraged that both reviewers recognised the value of the empirical contribution: the semantic advantage is robust across experiments, the two new experiments are pre-registered, and the drift-diffusion modelling provides an informative decomposition of behavioural performance. At the same time, both reviewers raise an important and convergent point: the manuscript currently places too much interpretive weight on non-decision time and sometimes moves too quickly from decision-model parameters to claims about the representational format of working memory.

We agree that this aspect of the manuscript should be revised. In the next version, we will substantially soften claims about adaptive reformatting and long-term-memory-like formats. We will instead frame the central contribution more precisely: semantic-category judgements show a reliable advantage at stages preceding evidence accumulation and this advantage is modulated by attentional prioritisation and interference during the maintenance interval. Our data constrain the dynamics with which different kinds of information become available for WM-guided decisions, but they do not, on their own, provide a direct measure of representational format. This hypothesis should be tested in future experiments.

At the same time, we think the data provide stronger constraints on alternative explanations than the current manuscript makes clear. The reviewers correctly note that non-decision time is not a pure retrieval parameter, as we also note in the discussion. It can include probe encoding, response preparation, motor execution, and other processes. We will therefore avoid more explicitly equating NDT directly with retrieval latency. However, many of the alternatives raised by the reviewers, such as easier question reading or simpler response mapping for semantic probes, predict a relatively fixed semantic-perceptual offset. In our experiments, the probes and response mappings are held constant across attentional conditions, while the semantic NDT advantage changes as a function of whether the relevant item can be prioritised in advance or must be selected/reactivated at test. We will restructure the Results and Discussion to make these condition $\times$ feature interactions central to the argument.

We will also clarify the logic of Experiment 1. We agree with Reviewer 1 that a valid retro-cue likely triggers retrieval or reactivation of the cued item. Our original phrasing, which described the valid-cue condition as reducing retrieval demands, was imprecise. The critical manipulation is better described as shifting item prioritisation/retrieval earlier in the trial. Under valid cueing, the relevant item can be prioritised before the probe appears, whereas under neutral cueing, item selection and access must occur after probe onset. We will rewrite this section accordingly.

We will also clarify our operationalisation of semantic and perceptual categories. The present contrast is specifically between semantic category information (animate versus inanimate) and perceptual-format information (photograph versus drawing). We agree that the perceptual judgement is still categorical and does not measure fine-grained perceptual fidelity. We will therefore avoid broad claims about semantic structure or perceptual detail in general. However, as pointed in the manuscript, we believe the contrast remains meaningful: the two dimensions are orthogonal within the same stimuli, and previous work using the same feature space showed the opposite ordering during perception (Linde-Domingo et al., 2019), where perceptual-format information was available before semantic-category information. We will move this argument earlier in the manuscript and present it as converging evidence for dissociable access dynamics, while acknowledging that it does not by itself prove representational format.

In response to the modelling concerns, we will expand the model-validation section. Specifically, we plan to add posterior predictive checks for the reported models, report model comparisons more transparently, clarify when more complex models do or do not provide practically meaningful improvements, and include sensitivity analyses using alternative parameterisations where identifiable.

We will also make several methodological clarifications. First, because the reanalysis of Kerrén et al. (2022) forms a substantial part of the manuscript, we will add a fuller description of the original task in the main text, including how the probed item was indicated at test. Second, we will rewrite the unclear sentence describing pseudo-random stimulus selection in Experiment 1 and add a control analysis testing whether performance differs when the probed item belongs to the majority versus minority category within the trial. Third, we will clarify the stimulus repetition scheme and discuss possible long-term-memory contributions. Importantly, because semantic and perceptual probes are applied to the same items from the same trials, any repetition history or proactive-interference contribution is shared across the two probe types, although we agree that this should be discussed explicitly.

Finally, we will revise the broader theoretical framing. We will remove or substantially qualify claims linking the present data directly to episodic memory and imagery. We will also integrate the recent literature suggested by Reviewer 2 on semantic structure, associative relations, long-term-memory contributions to working memory, and boundary conditions for semantic labelling effects. This will allow us to position the study as one piece of a broader literature on how semantic information influences WM performance, rather than as direct evidence for a general representational reformatting mechanism.

In summary, the revised manuscript will make a narrower but stronger claim: semantic-category information shows a robust pre-accumulation advantage during WM-guided decisions, and this advantage is shaped by attentional prioritisation and interference during maintenance. We will present this as evidence about WM access dynamics and decision components, not as direct evidence that WM representations are transformed into long-term-memory-like formats.

<https://doi.org/10.7554/eLife.111859.1.sa0>