Article:
Atkinson, A.L., Allen, R.J., Waterman, A.H. et al. (2026) How do individuals direct attention to valuable information in working memory tasks? Evidence from eye-tracking. Journal of Memory and Language, 150. 104781. ISSN: 0749-596X

https://doi.org/10.1016/j.jml.2026.104781

How do individuals direct attention to valuable information in working memory tasks? Evidence from eye-tracking

Amy L. Atkinson[1], Richard J. Allen[2], Amanda H. Waterman[2], Naziya Lokat[1], Inez Sharp[1], Amie Suthers[1], Tom Beesley[1]

1. Department of Psychology, Lancaster University, Bailrigg, Lancaster, UK.

2. School of Psychology, University of Leeds, Leeds, UK.

*Corresponding author; amy.atkinson@lancaster.ac.uk.

Status: This is the author accepted manuscript. This paper has been accepted for publication in *Journal of Memory and Language*. The DOI of the published article is: 10.1016/j.jml.2026.104781.

Open science practices: The data and analysis scripts are publicly available at: https://osf.io/4fk9m

Conflicts of interest: The authors report no conflicts of interest.

Funding: This work was funded by an Experimental Psychology Society small grant to ALA.

Ethics statement: All procedures were performed in compliance with relevant laws and institutional guidelines. The research received ethical approval from the Faculty of Science and Technology Ethics Committee at Lancaster University (FST-2022-0660). The privacy rights of human subjects have been observed and informed consent was obtained from participants prior to the study.

CRediT statement: ALA – Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Visualization; Writing – original draft. RJA – Conceptualization, Methodology, Writing – review & editing. AHW - Conceptualization, Methodology, Writing – review & editing. NL – Investigation, Methodology, Writing – review & editing. IS - Investigation; Methodology; Writing– review & editing. AS - Investigation; Methodology; Writing– review & editing. TB – Conceptualization, Formal analysis, Methodology, Project administration, Software, Supervision, Visualization, Writing – review & editing.

Abstract

There is robust evidence that individuals can direct their attention towards valuable information in working memory. However, the cognitive mechanisms underlying such effects are still under debate. Whilst attentional refreshing is considered to at least partially drive the effect, research has yet to establish whether other processes also contribute. One possibility is that participants preferentially encode high value information, creating a stronger representation relative to other information presented. We tested this hypothesis in two experiments. Participants completed a visual working memory task, which involved remembering the color of shapes for a brief period of time. During the encoding phase participants' eye movements were recorded to assess fixation behaviors. Participants were then presented with one shape and asked to report the color as accurately as possible on a color wheel. In Experiment 1, participants experienced trials where one particular item was more valuable (i.e., worth more notional points than the rest), or trials where all items were of equal value. In Experiment 2, each item within a trial differed in its value in a graded manner (i.e., 1-2-3-4 points). Across both experiments, participants exhibited the lowest recall error for the highest value information. Moreover, participants spent significantly longer fixating high value items during the encoding phase relative to other items, and often encoded high value items first. These eye movement behaviors partially mediated the effect of value on recall error. Taken together, these experiments rule out the possibility that maintenance mechanisms solely drive value-based prioritization effects, demonstrating an important role of attention during encoding.

Keywords: working memory, value-based prioritization, value-driven prioritization, strategic prioritization, reward, attention, eye movements

Highlights:

1. Participants encode high value information more strongly than other information
2. High value items are often fixated on first (resulting in a "primacy-like" effect)
3. Participants fixate longer on high value items during encoding
4. Encoding strategies partly drive value-based prioritization effects in working memory
5. Value-based prioritization cannot solely be explained by maintenance mechanisms

How do individuals direct attention to particularly valuable information in working memory tasks? Evidence from eye-tracking

Working memory (WM) is a system that temporarily stores and processes a limited amount of information for a brief period of time (Baddeley et al., 2021; Cowan, 2017). It is considered crucial for a wide range of cognitive activities, including decision-making (Cui et al., 2015), reasoning (Kyllonen and Christal, 1990; Süß et al., 2002), and learning (Cowan, 2014). As the capacity of working memory is extremely limited, the ability to focus attention on particularly important or goal-relevant information is vital (Allen et al., 2025; Oberauer & Hein, 2012). Indeed, a close relationship between working memory and attention has been demonstrated in the literature (e.g. Allen et al., 2017; 2025; Awh & Jonides, 2001; Oberauer, 2019; Vogel et al., 2005), with attention forming a key aspect of most prominent working memory models (e.g. Allen et al., 2026; Baddeley et al., 2021; Barrouillet & Camos, 2021; Cowan, 1999; Cowan et al., 2021; Engle, 2002; Hitch et al., 2025; Oberauer, 2002; 2013).

A large body of research has now demonstrated that individuals can direct their attention to particular items in working memory (see Allen et al., 2025 and Souza & Oberauer, 2016 for reviews), thereby ensuring they remain at least briefly accessible in conscious awareness (Allen et al., 2026; Cowan et al., 2025). Studies examining this ability have often used visual cues (e.g. arrows), which inform participants which item will, or is more likely to, be tested at retrieval (Souza & Oberauer, 2016). In recent years, value-based prioritization has also been increasingly used to explore the ability to direct attention in working memory (e.g. Allen & Ueno, 2018; Atkinson et al., 2018; 2021; 2022; 2024; Hautekiet et al., 2024; Hitch et al., 2018; Hu et al., 2014; 2016; Sandry et al., 2014; Sandry & Ricker, 2020; Vergauwe et al., 2025). In this paradigm, participants are presented with items to remember for a short period of time. Typically, prior to encoding, participants are told that one (or more) items are relatively more valuable than the others. Performance is then compared to a condition in which another item is more valuable (e.g. Hu et al., 2014; 2016), or a condition in which all items presented are of equal value (e.g. Atkinson et al., 2018; 2021; 2022; 2024; 2025). Value is typically manipulated by attributing notional points to the items. In such tasks, participants typically show superior memory for high value items relative to low value items (e.g. Atkinson et al., 2022; Hu et al., 2014; 2016), or a condition in which all items are of equal value (e.g. Atkinson et al., 2018; 2021; 2022; 2024; 2025). Prioritization boosts for high value information does, however, typically come at a cost to low value information, which is not remembered as accurately (Allen et al., 2025; Atkinson et al., 2021; 2022; 2025; Vergauwe et al., 2025). As such, prioritization does not appear to increase overall working memory capacity, with effects instead thought to emerge from a re-allocation of attention towards high value item(s) and away from less valuable items (Allen et al., 2025; Atkinson et al., 2022; Hitch et al., 2020).

Value-based prioritization effects in working memory have been demonstrated across a wide variety of task contexts. For example, existing studies have demonstrated that individuals can prioritize valuable information across a range of presentation methods (e.g. simultaneous and sequential presentation; e.g. Allen & Ueno, 2018; Atkinson et al., 2019; 2022; Hu et al., 2014; 2016) and modalities (e.g. visual, verbal, tactile, and cross-modal representations; Atkinson et al., 2018; 2021; Cinar et al., 2025; Roe et al., 2024; Sandry et al., 2014). Moreover, individuals can direct their attention in a highly graded manner when all items differ in their value (i.e. point values of 1-2-3-4 are used) (Allen & Ueno, 2018), and are also able to prioritize multiple items concurrently (Allen & Ueno, 2018; Hitch et al., 2018). Finally, whilst most research has studied young non-clinical adult samples, value-based prioritization effects have also been demonstrated in a number of other groups, including healthy older adults (Allen et al., 2021), children (Atkinson et al., 2019), and individuals with symptoms of ADHD (Atkinson et al., 2025).

Although there is robust evidence that individuals can prioritize valuable information in working memory (Allen et al., 2025), the cognitive mechanisms underlying such effects are not well understood. One potential mechanism is verbal rehearsal. However, within visual tasks (which have most commonly been used to examine value-based prioritization; Allen et al., 2025), articulatory suppression is commonly implemented to disrupt participants' ability to verbally recode and rehearse information (e.g. Atkinson et al., 2018; Hu et al., 2014; 2016). Despite this, large prioritization effects are typically observed (Allen et al., 2025), suggesting verbal rehearsal is unlikely to play a critical role. Indeed, the role of verbal rehearsal has been considered more directly in studies which have compared value-based prioritization effects across conditions which do, and do not, require articulatory suppression. Using visually presented verbal information (letters), Sandry et al. (2014) found no significant difference in the size of prioritization effects when participants did, and did not, engage in articulatory suppression (Sandry et al., 2014). Further, Atkinson et al. (2021) found that value-based prioritization effects were *larger* in size when participants engaged in articulatory suppression during a task requiring retention of auditorily-presented digits. Thus, it appears unlikely that verbal rehearsal is the main mechanism driving value-based prioritization effects in working memory.

Although the cognitive mechanisms involved in value-based prioritization were not directly examined, a study by Hu et al. (2016) provides relevant insights. Hu et al. (2016) demonstrated that value-based prioritization effects were either reduced in size or abolished when participants engaged in an attentionally-demanding concurrent task during encoding and maintenance. This suggests that value-based prioritization effects in working memory are reliant on an attentionally-demanding process that occurs during encoding, maintenance, or across both phases. Based on such findings, several potential mechanisms have been proposed to explain value-based prioritization effects (Hu et al., 2016). One such mechanism is attentional refreshing (Allen et al., 2025; Atkinson et al., 2021; Hu et al., 2016; Sandry et al., 2014). Attentional refreshing is a domain-general mechanism that involves participants "thinking of" particular items during the retention interval (Camos et al., 2018). This maintenance mechanism is separable from verbal rehearsal (Camos et al., 2009; Camos & Barrouillet, 2014; Mora & Camos, 2015) and is considered to be crucially reliant on attention (Camos & Barrouillet, 2014; Camos et al., 2018). It has been suggested that when one item is more valuable, participants may adopt a biased attentional refreshing approach, refreshing this item more frequently and/or for longer periods of time relative to other to-be-remembered information (Allen et al., 2025; Atkinson et al., 2021; Hu et al., 2016; Sandry et al., 2014).

This hypothesis was tested by Atkinson et al. (2022), by combining the value-based prioritization task with a directed-refreshing paradigm developed by Souza and colleagues (Souza et al., 2015; 2018; Loaiza & Souza, 2022; Vergauwe et al., 2024). Participants were presented with arrays of four colored shapes, to be remembered for a brief period. Participants were then presented with one of the shape outlines and asked to report the color it had previously been presented in using a color wheel. As in the typical value-based prioritization paradigm, participants were told the point values associated with each item prior to encoding. In some trials, all items were equally valuable. In other trials, one of the items was more valuable (i.e. worth more notional points) than the rest. During maintenance, the screen either remained blank or arrow cues were presented which pointed to the spatial location in which items had appeared during encoding. In trials where arrow cues were presented, the to-be-tested item was either cued or not cued (uncued). In trials where cues were presented during maintenance, participants were told to think of the cued item for the entire duration the cue was presented. If participants were already preferentially refreshing the high value item, they should receive no additional boost when instructed to refresh this item, whereas they should receive a boost for equal and low value items. Alternatively, if a biased attentional refreshing process does not drive the value boost, instructions to refresh should enhance performance of all items to a similar extent. Support for the former was observed, with instructions to refresh enhancing memory for the equal and low value items, but not the high value items. This was taken as evidence that a biased attentional refreshing approach at least partially explains participants' ability to prioritize valuable information in working memory (Atkinson et al., 2022).

However, it is possible (and indeed likely) that other mechanisms are also involved in generating value-based prioritization effects in working memory (e.g. Allen & Atkinson, 2021; Allen et al., 2025). As the point values are typically presented in advance of encoding, one possibility is that participants encode the high value item with greater strength relative to other items in the array or sequence (Allen et al., 2025). For example, when items are presented simultaneously, participants may preferentially encode the high value item by spending longer fixating on this item during encoding. This would likely result in better memory for high value items, as previous research has suggested that memory performance improves as the proportion of encoding resource assigned to an item increases (Bays & Husain, 2008). Existing studies do provide some indirect evidence to suggest an important role of encoding processes. For instance, although large value-based prioritization effects are typically observed when value information is presented in advance of encoding (see Allen et al., 2025), effects are not consistently observed when value information is presented after encoding (Allen & Atkinson, 2021; Jeanneret et al., 2023; Hautekiet et al., 2024). This may suggest that behaviors that participants engage in during the encoding phase are an important driver of value-based prioritization effects. However, it is important to note that prospective and retrospective prioritization methods are not entirely comparable, with smaller retrospective prioritization effects potentially reflecting items being less accessible when the value information becomes available (Allen & Atkinson, 2021; Allen et al., 2025). Thus, although there is some evidence that encoding processes may be important in generating value-based prioritization effects, the empirical evidence thus far provides only indirect support for this hypothesis. Further research is therefore needed to more directly examine this possibility. Such findings would provide important theoretical insights into the cognitive mechanisms underlying value-based prioritization effects (Atkinson et al., 2022; Allen et al., 2025), as well as providing important insights into how individuals direct attention in working memory more generally.

A pair of experiments were therefore conducted to examine how participants direct attention in working memory in a value-based prioritization task, by examining eye movements during the encoding phase. In recent years, a number of studies have utilized eye-tracking to investigate working memory processes (e.g. Bhanap et al., 2025; Loaiza & Souza, 2022; Morey et al., 2018; Nikolov et al., 2025; Souza et al., 2020; Ueno & Allen, 2025). However, there have not yet been any studies utilizing eye-tracking to understand value-based prioritization effects in working memory. Moreover, very little research has utilized eye-tracking to examine how individuals direct attention in working memory more generally (Matsukura et al., 2007; Ueno & Allen, 2025). Experiment 1 explored how participants' eye movements vary as a function of value, and whether eye movements drive value-based prioritization effects. Experiment 2 then examined whether eye movements are involved in generating value-based prioritization effects when value is manipulated in a more fine-grained manner (i.e., each item in the array is worth a different number of points, ranging from 1-4; Allen & Ueno, 2018).

Experiment 1

Participants completed a visual working memory task, which involved remembering arrays of four unique shapes presented simultaneously for a brief period of time. The outline of one shape was then presented at test, and participants were asked to reproduce the color this item had previously been presented in using a color wheel. This approach is similar to that taken by Atkinson et al. (2022), who investigated the extent to which biased attentional refreshing can account for value-based prioritization effects in working memory. Prior to encoding, participants were presented with numbers in the locations of the four to-be-presented items, which corresponded to the number of points participants would obtain if they were asked about the item presented at that spatial location and they responded with a sufficient level of accuracy (Allen & Ueno, 2018; Atkinson et al., 2022). Across trials, the items could be subjected to one of three point value conditions: "High", "Low", and "Equal". In some trials, all items were equally valuable (i.e. worth the same number of notional points; "Equal"). In other trials, one of the items was more valuable (i.e. worth more notional points; “High”) than the rest (“Low”). Eye movements were recorded throughout the trials, though we were particularly interested in behaviors during the encoding phase.

In line with previous studies, it was expected that a value-based prioritization effect would emerge, with participants showing the lowest recall error for high value information, intermediate for equal value information, and the highest recall error for low value information (e.g., Atkinson et al., 2022). Of particular interest was whether participants would spend a greater amount of the encoding phase fixating high value items relative to equal or low value items, and whether such patterns would drive the effects of value on task performance. Moreover, we examined the extent to which participants fixated the high value item first, and whether this also acted as a mediator.

Method

Participants

A power calculation was conducted in G*Power (Faul et al., 2007) for a repeated measures one-way ANOVA with three value levels (High, Equal, Low). Based on a medium effect size ($f = 0.25$) and alpha set to 0.05, a sample of 36 participants was required to provide 90% statistical power. Forty-one participants completed the experiment. All participants confirmed they were between 18-35 years of age, had normal or corrected-to-normal vision, no color blindness, English as a first language, and no literacy difficulties (e.g., dyslexia) or neurological conditions. Thirty-six participants provided demographic information (*M.age* = 22.39, *SD* = 4.10; Gender: 29 women, 6 men, 1 non-binary). One participant was excluded from the study due to a computer failure. A further two participants were removed from all analyses, as the percentage of trials with no recorded fixations during the encoding phase of the task was 2.5 standard deviations above the mean. The final analysis was therefore performed on data from 38 participants, of which 33 provided demographic data (*M.age* = 22.33, *SD* = 4.17; Gender: 26 women, 6 men, 1 non-binary). Ethical approval was granted by the Faculty of Science and Technology Ethics Committee at Lancaster University (Ref. number: FST-2022-0660).

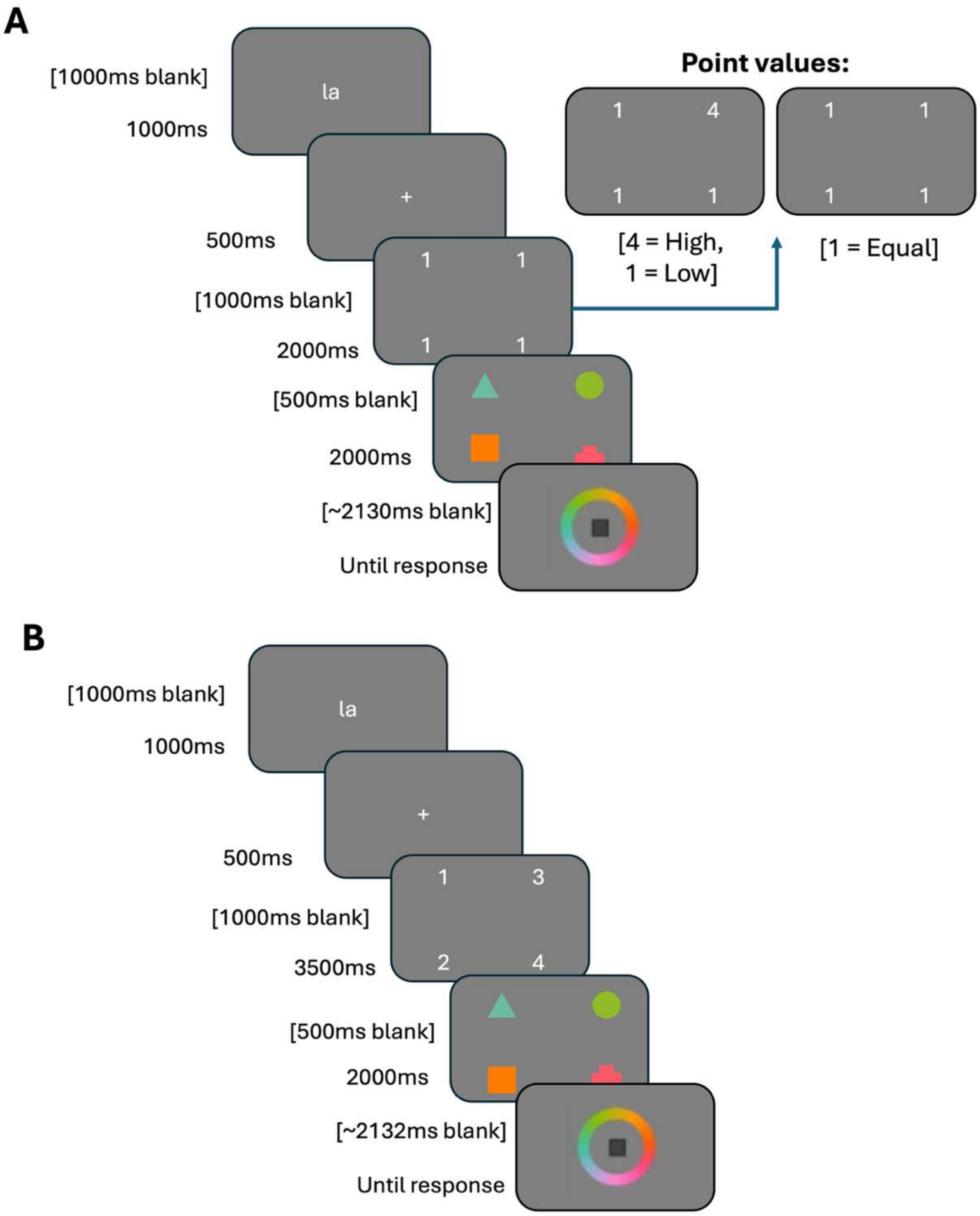
Design, Materials, and Procedure

Value was the independent variable (within-subject), with three levels (High, Equal, Low). The main dependent variables were recall error and the proportion of time participants spent fixating on items during the encoding phase.

The experiment was run using PsychoPy 3 (Peirce et al., 2019) on a desktop computer (screen resolution = 1920 x 1080 pixels), with eye movements recorded using a Tobii Pro X3-120 Eye Tracker (sampling at 120hz). Participants completed one in-person session. Participants first completed a 6-point calibration. They then completed the memory task. The experimental paradigm used is displayed in Figure 1A. Participants were first presented with a blank screen for 1000ms, followed by the word "la" for 1000ms. They were asked to repeat this aloud until the test phase to disrupt verbal recoding. Compliance was monitored by a research assistant who remained present throughout the testing session. Participants then saw a fixation cross for 500ms, followed by a blank screen for 1000ms. Next, the point values were presented for 2000ms. After a blank delay of 500ms, the four colored shapes were presented for 2000ms. In each array, a square, circle, triangle and cross were presented simultaneously (each measuring approximately 1.3° visual angle, based on a viewing distance of 50cm). Each shape was presented at one of four spatial locations, positioned at the corners of an imaginary square measuring approximately 22x22° visual angle located at the center of the screen, based on a viewing distance of 50cm. The colors of the shapes were randomly selected from 360 values, evenly spaced around a circle in the CIELAB color space (L = 70, a = 20, b = 38, radius = 60), with the constraint that colors presented within each trial could not be within 40° of each other. There was then a blank retention interval. The duration of this retention interval differed marginally between trials due to very slight variability in the time taken by the experimental software to draw the color wheel that was presented at test (*M.* across participants = 2130ms, *SD* across participants = 7ms, Range across participants = 2116-2145ms). The test phase then commenced, with the presentation of one of the shapes at the center of the screen (measuring approximately 1.3° visual angle, based on a viewing distance of 50cm). Each of the shapes were equally likely to be probed. The probed item was surrounded by a color wheel, presented as an annulus with inner radius of 25% of the screen height and an outer radius of 33% of the screen height, with a random rotation on every trial. Participants were asked to reproduce the color of the probed shape as accurately as possible by clicking on the color wheel using a mouse. There was no time limit. As in previous studies (see Allen et al., 2025), feedback was not provided. Eye movements were recorded throughout the trial. Eye movements were recorded from both eyes and averaged. For samples where there was data from only a single eye, that eye was used.

Figure 1

The Experimental Paradigm Used in Experiment 1 (Panel A) and Experiment 2 (Panel B)



Note. The number (1 or 4 in Experiment 1, 1-4 in Experiment 2) reflects the point values associated with the item appearing at that spatial location. Figure not to scale.

As part of the instructions, participants were told that the numbers appearing before the items indicated the number of points participants would collect if they were asked about that item and they responded accurately (i.e. "1" indicates the item appearing at that spatial location was worth 1 point, "4" indicates the item appearing at that spatial location was worth 4 points). Participants were explicitly informed that the point values would not predict which item would be tested. As in Atkinson et al. (2022), the precise level of accuracy required to collect the points was not specified to ensure participants responded as accurately as possible. In the equal value arrays, all numbers in the array were "1", indicating that all items were worth one point. In the differential value arrays, one item was worth four points (indicated by the number "4"), whilst the other items were worth one point (indicated by the number "1"). This resulted in three value conditions: (i) "Equal" value - where all items in the array were worth the same number of points; (ii) "High" value - items worth 4 points in arrays where one item differed in value relative to the rest; and (iii) "Low" value – items worth 1 point in arrays where one item differed in value relative to the rest. In the trials where items differed in their value, the spatial location of the high value item was counterbalanced. As in previous studies (see Allen et al., 2025), points were notional and not linked to a meaningful reward (e.g., money).

Participants completed 220 test trials in total. There were 44 test trials where all items were equally valuable, and 176 test trials where one item differed in its value relative to the rest. As each item in the array was equally likely to be tested, the High value item was tested on 44 test trials (out of 176 where one of the items differed in its value) and a Low value item was tested on 132 test trials (out of 176 where one of the items differed in its value). Thus, there were an equivalent number of trials where High and Equal value items were tested, and necessarily many more trials in which a Low value item was tested. All trial types were intermixed. Each of the four spatial locations were tested an equal number of times within each value condition. Participants completed five practice trials at the start of the experiment, which were not analyzed.

Data Analysis

Data and analysis scripts for both experiments are available at https://osf.io/4fk9m. All analyses were conducted in R.

Analyses were first conducted on the task outcomes. The main dependent variable in this analysis was recall error, calculated as the deviation between the correct color and the color selected, which ranged between −180° to 180°. Recall error was computed as the absolute value of this deviation. Frequentist analyses were conducted using the *afex* (Singmann et al., 2023) and *emmeans* (Lenth, 2023) packages. Where necessary, Greenhouse-Geisser (GG) corrections were made. Outcomes from post-hoc tests were corrected for multiple comparisons using the Bonferroni-Holm method. Bayes Factor (BF) analyses were also conducted using the *BayesFactor* package (Morey & Rouder, 2022). BF analysis indicates the strength of evidence for the presence or absence of an effect. Bayesian ANOVAs were conducted using default priors, with the number of iterations set to 500,000 and the parameter *which_model* set to "top". Reported BF_{10} values indicate the support for the alternate model over the null model, where values over 3 are typically interpreted as providing substantial evidence in support of that model, and values below 1/3 are taken as greater support for the null model. Our conclusions are primary based on the frequentist outcomes, but we draw readers attention to instances where interpretation would differ if conclusions were based on the BF analyses. Behavioral analyses were conducted using data for all trials for participants included in the final sample ($N = 38$).[1] The supplementary materials report analyses of response times to ensure that any effect of value on recall error was not due to a speed-accuracy trade off, whereby participants might respond more precisely because they took longer to give their answer. However, there was no evidence of a speed-accuracy trade off in either experiment.

The data were fit using a Bayesian implementation of the three-component mixture model (Bays et al., 2009) using the *bmm* package (Frischkorn & Popov, 2025) to establish whether any value effects observed reflect an increased probability of recalling the target item, increased precision, decreased probability of recalling a non-target item, or decreased guessing. The model was fit using four chains, each with 2,500 warm up samples (which were discarded), and 2,500 samples that were used, giving a total of 10,000 samples. Chain convergence was assessed using the Gelman–Rubin $\hat{R}$ statistic (Gelman & Rubin, 1992). This

[1] Behavioral analyses were also conducted excluding trials with no recorded fixations during the encoding phase (to match the eye-tracking analyses described below), as well as including the two participants who were excluded from analyses due to poor tracking ($N = 40$; see supplementary materials).

indicated good convergence (all $\hat{R}$ values = 1.00). Differences in posterior distributions were computed for comparisons of interest. Differences were judged as credible if the 95% highest density interval did not cross zero. The "probability of direction" (*pd*) statistic was also computed using bayestestR (Makowski et al., 2019). This indicates the proportion of the posterior distribution that is in the same direction as the median (minimum = 0.5, maximum = 1; Kelter, 2020). As with the main behavioral analyses, modelling was run using data for all trials from participants included in the final sample (N = 38).

Next, we analyzed the eye-tracking data. Eye data were processed using the *eyetools* package (Beesley & Ivory, 2025). Fixations were identified with the maximum tolerance for dispersion (in pixels) set to 75. The minimum duration of fixations was set at the default of 150ms. Areas of interest were 300x300 pixels around the center of each to-be-remembered item. In the final sample (N = 38), the mean percentage of missing samples during the encoding phase was 1.32% (SD = 1.70), whilst the mean percentage of trials that had no recorded fixations during this phase was 0.80% (SD = 2.56). Trials were excluded where there were no detected fixations during the encoding phase. Following this, the mean number of fixations per trial in the encoding period was 4.93 (SD = 0.75, Range = 3.14-6.36).

The main dependent variable of interest in the eye-tracking analyses was the total proportion of time participants spent fixating on an item during encoding. Both Frequentist and Bayesian analyses were conducted, as described above. Mediation analysis was conducted to investigate whether the relationship between value and recall error was mediated by the proportion of time participants spent fixating on the tested item during the encoding phase. This was conducted separately for each of the value pair comparisons: the overall value effect (Low vs High value), the value boost (Equal vs High value), and the value cost (Equal vs Low value). Indirect effects and the proportion mediation were computed using the *mediation* package (Tingley et al., 2014), with the number of simulations set to 5,000. All models included a random intercept of participant ID.

Analyses were then conducted to investigate whether the proportion of trials in which participants fixated on the High value item first was greater than expected based on an equal distribution (0.25). This factor was also explored as a mediator using the same approach outlined above.

A number of subsidiary analyses were then conducted. Evidence that the proportion of time participants spent fixating on an item during encoding differed as a function of value could reflect: (i) participants fixating on higher value items more frequently; (ii) participants having longer fixations towards higher value items; or (iii) a combination of the above. Thus, where there was evidence that the total proportion of time participants spent fixating on an item during the encoding phase differed as a function of value, further analyses were conducted to delineate between these possibilities. These analyses are reported in the supplementary materials.

For completeness, we also analyzed whether the total proportion of time participants spent fixating on an item during maintenance (the period following encoding where no stimuli were present in the display) differed as a function of the value of items that had appeared in those locations. We also examined this factor as a potential mediator. The outcomes are presented in the supplementary materials.

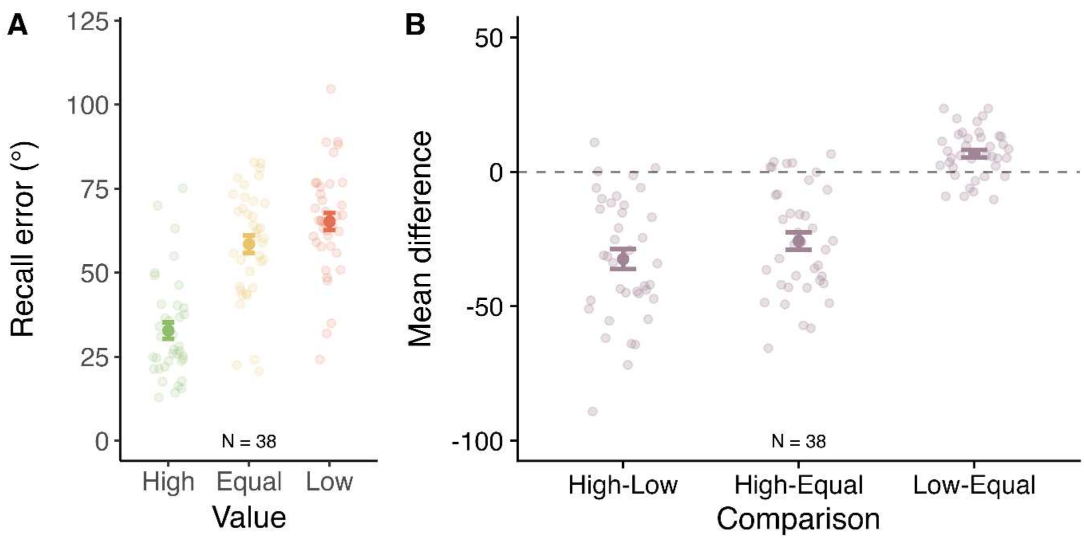
Results

Recall Error

Mean recall error is displayed in Figure 2A as a function of value. Meanwhile, Figure 2B displays the difference in recall error between the three value conditions. A one-way repeated measures ANOVA on the factor of value (High, Equal, Low) revealed a significant effect of value (GG-corrected, $F(1.23, 45.37) = 65.61$, $MSE = 277.36$, $p < .001$, $\eta_p^2 = .64$; $BF_{10} >$ 10,000). Post-hoc tests revealed significantly lower recall error for High value items (M = 32.8, SE = 2.43) relative to Equal value (M = 58.5, SE = 2.61; $p < .001$, d = -1.28; $BF_{10} >$ 10,000) and Low value (M = 65.2, SE = 2.59; $p < .001$, d = -1.40; $BF_{10} >$ 10,000) items. Recall error was also significantly lower for equal value items relative to low value items ($p < .001$, d = -0.75; BF_{10} = 474.33).

Figure 2

Task Performance in Experiment 1



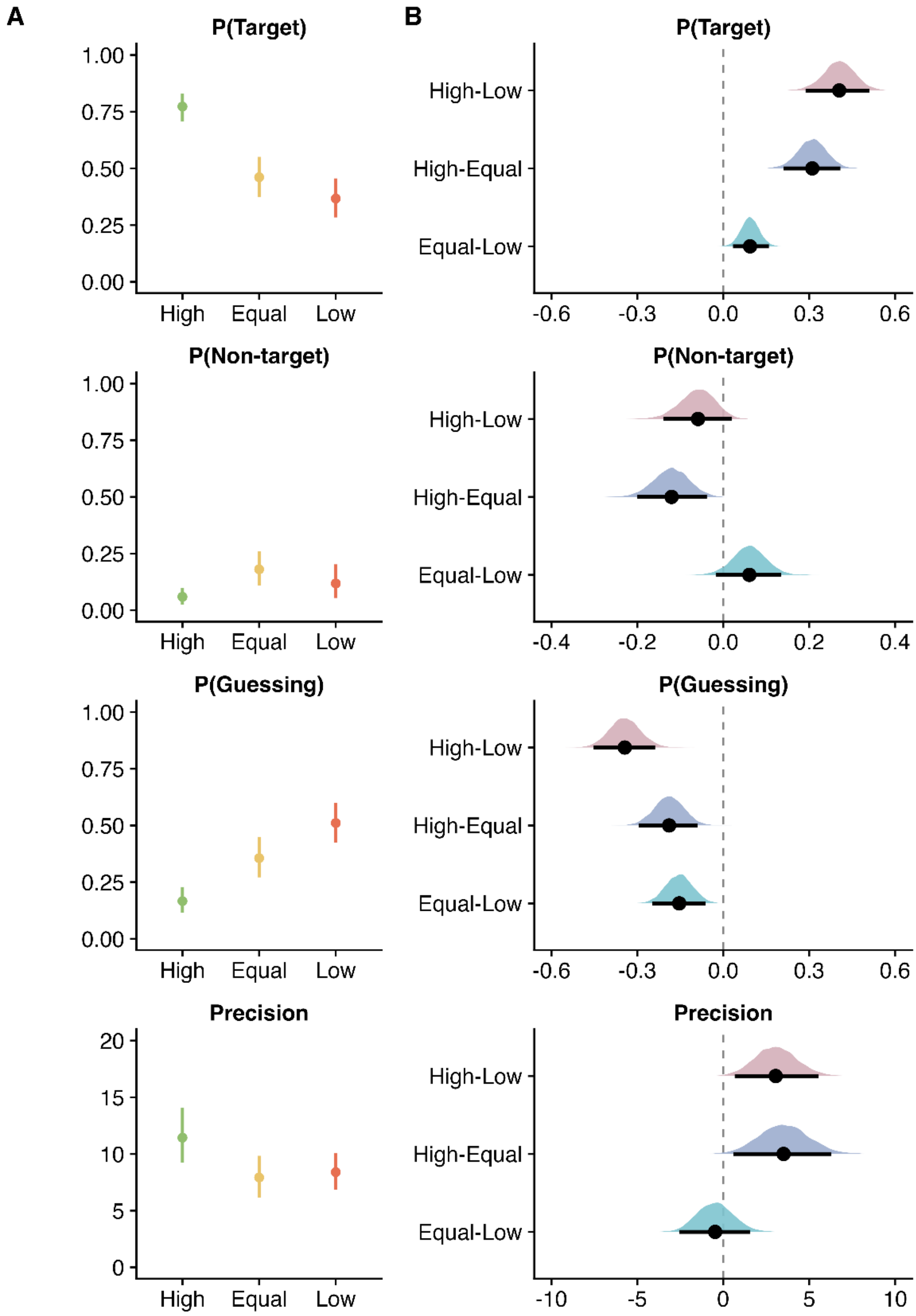
Note. Panel A displays mean recall error, as a function of value (lower recall error reflects better performance). Panel B displays the mean difference between the three value conditions. The dotted line at 0 in Panel B reflects no difference. Values above 0 indicate higher recall error (i.e. worse performance) in the first condition listed, whilst values below 0 indicate lower recall error (i.e. better performance) in the first condition listed. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error (SE) of the mean.

Bayesian Mixture Modelling

The parameter estimates from the hierarchical Bayesian mixture model are displayed in Figure 3A, as a function of value. The differences in posterior distributions were calculated between all three value conditions (High vs Low, High vs Equal, Equal vs Low). The median, 95% highest density intervals, and distributions are presented in Figure 3B. The modelling revealed that the probability of recalling the target item was higher in the High value condition, relative to Equal and Low value conditions (*pd* = 1.000 for both comparisons). The probability of recalling the target item was also higher in the Equal value condition relative to the Low value condition (*pd* = .998). The probability of recalling a non-target item was lower in the High value condition than the Equal value condition (*pd* = .999), with no credible differences between the other conditions (High vs Low, *pd* = .940; Equal vs Low, *pd* = .938). Furthermore, the probability of guessing was lower in the High value condition than the Equal and Low value conditions (*pd* = 1.000 for both comparisons), and was also lower in the Equal value condition than the Low value condition (*pd* = .999). Finally, precision was higher in the High value condition than the Equal (*pd* = .993) and Low value conditions (*pd* = .996), with no credible difference between the Equal and Low value conditions (*pd* = .677).

Figure 3

Outcomes from the Bayesian Mixture Modelling in Experiment 1



Note. Panel A: Estimates from the hierarchical Bayesian mixture modelling as a function of value. Points reflect the median, whilst error bars display 95% highest density intervals. Panel B: Difference between the posterior distributions for the parameters in the Bayesian mixture modelling. For each parameter, differences are displayed for all comparisons between the value conditions (High vs Low, High vs Equal, Equal vs Low). The black circle displays the median, the black horizontal bar shows the 95% highest density intervals, and the color shading reflects the distribution. The grey vertical line at 0 reflects no difference. In both panels, facets from top to bottom show: probability of recalling the target item, probability of recalling a non-target item, probability of guessing, and precision.

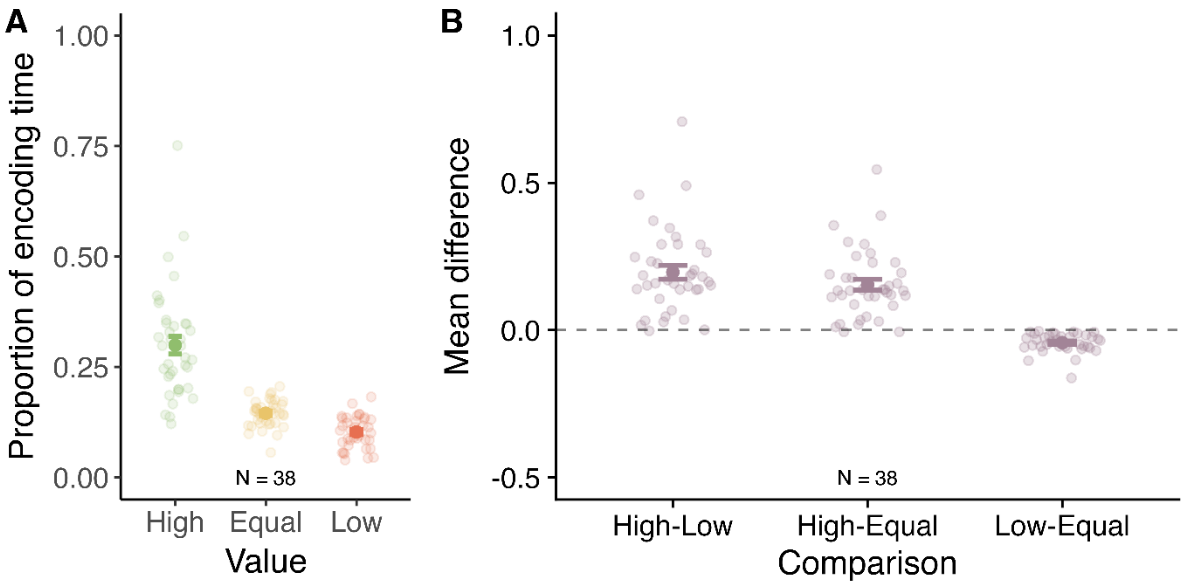
Eye-Tracking Outcomes

Encoding Fixation Duration. Figure 4A displays the proportion of encoding time participants spent fixating items, as a function of item value. Figure 4B shows the differences in these proportions between value conditions. A one-way repeated measures ANOVA on the factor of value (High, Equal, Low) found a significant effect of value (*GG-corrected* $F(1.01, 37.43) = 68.46$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .65$; $BF_{10} > 10{,}000$). Participants spent a greater proportion of the encoding phase fixating High value items ($M = 0.30$, $SE = 0.02$) relative to the Equal ($M = 0.15$, $SE = 0.01$; $p < .001$; $d = 1.33$; $BF_{10} > 10{,}000$) and Low value items ($M = 0.10$, $SE = 0.01$; $p < .001$, $d = 1.35$; $BF_{10} > 10{,}000$). Participants also spent longer fixating Equal value items relative to Low value items ($p < .001$; $d = 1.32$; $BF_{10} > 10{,}000$).[2]

Figure 4

The Proportion of the Encoding Phase Participants Spent Fixating on Items in Experiment 1

[2] Further subsidiary analysis was then conducted to establish whether this effect was driven by relatively more fixations towards High value items, individual fixations towards High value items being longer, or a combination (see supplementary materials for full reporting). To summarize, participants fixated on High value items more often than expected based on an equal distribution of fixations (0.25). Individual fixations to High value items were also longer relative to Equal and Low value items on average. Thus, participants fixate on High value more often and for longer than lower value items.

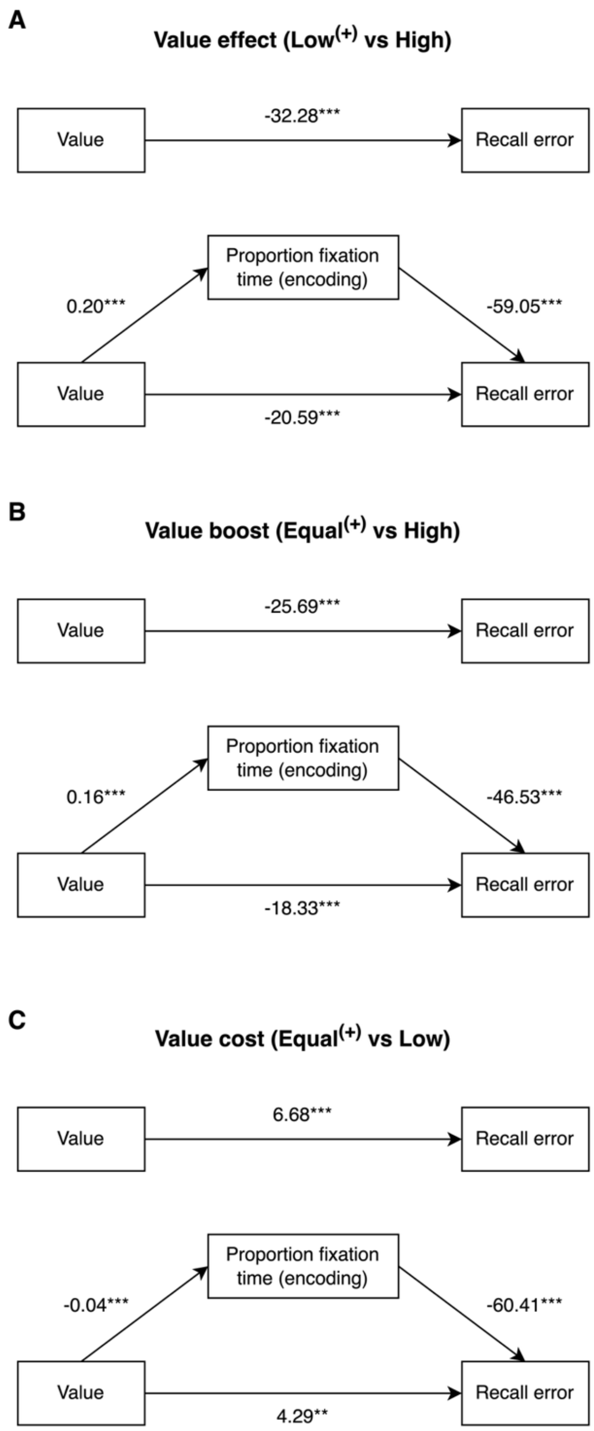


Note. Panel A displays the mean proportion of the encoding time participants spent fixating on an item as a function of item value. Panel B displays the mean difference in proportion of encoding time between the value conditions. The dotted line at 0 reflects no difference. Values above 0 indicate that participants fixated on the first condition listed for a higher proportion of the encoding phase relative to the second condition listed. Meanwhile, values below 0 indicate that participants fixated on the first condition listed for a lower proportion of the encoding phase relative to the second level listed. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error of the mean.

Mediation analysis was conducted to investigate whether the proportion of time participants fixated on the tested item during encoding mediated the relationship between value and recall error. As with the main eye-tracking analyses, this included only trials that had any fixations detected during the encoding phase. Separate mediation models were conducted for the overall value effect (Low vs High value), the value boost (Equal vs High value), and the value cost (Equal vs Low value). Outcomes are displayed in Figure 5. In all models, the effect of value on recall error was partially mediated by the proportion of time participants fixated on the tested item during encoding. The indirect effects were -11.70 for the overall value effect (Low vs High; 95% CI: -13.50 - -9.96; $p < .001$), -7.36 for the value boost (Equal vs High; 95% CI: -9.04 - -5.79; $p < .001$), and 2.40 for the value cost (Equal vs Low; 95% CI: 1.83-3.00; $p < .001$). The proportion of the effect mediated was 0.36 (95% CI: 0.30-0.43; $p < .001$) for the overall value effect (Low vs High), 0.29 (95% CI: 0.22-0.36; $p < .001$) for the value boost (Equal vs High), and 0.36 (95% CI: 0.23-0.65; $p < .001$) for the value cost (Equal vs Low). Taken together, these outcomes indicate the effect of value on recall error is partially, but not fully, mediated by the proportion of time participants spend fixating on an item during encoding.

Figure 5

Outcomes from the Mediation Analysis in Experiment 1, Examining Proportion of Fixation Time during Encoding as the Mediator

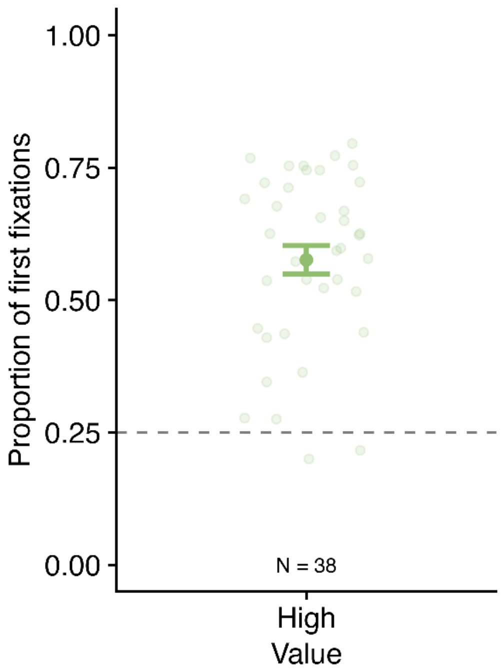


Note. Panel A reflects the overall value effect (Low vs High value), Panel B reflects the value boost (Equal vs High value), and Panel C reflects the value cost (Equl vs Low value). In all panels, the upper diagram reflects the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. $^{+}$ denotes reference category. *** $p < .001$, ** $p < .01$, * $p < .05$.

First Item Fixated. We then assessed whether participants were more likely to fixate a high value item first (see Figure 6). Trials were excluded where participants did not fixate on any of the items during the encoding phase. The mean proportion of trials in which the first item participants fixated was the High value item was 0.58 ($SE = .03$), which was significantly greater than that expected based on an equal distribution across items (0.25), $t(37) = 12.07$, $p < .001$, $d = 1.96$, $BF_{10} > 10{,}000$.

Figure 6

Mean Proportion of Trials in which the First Item Fixated was the High Value Item



Note. The dotted line at 0.25 reflects an equal distribution of first fixations across items. The bold point reflects the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error of the mean.

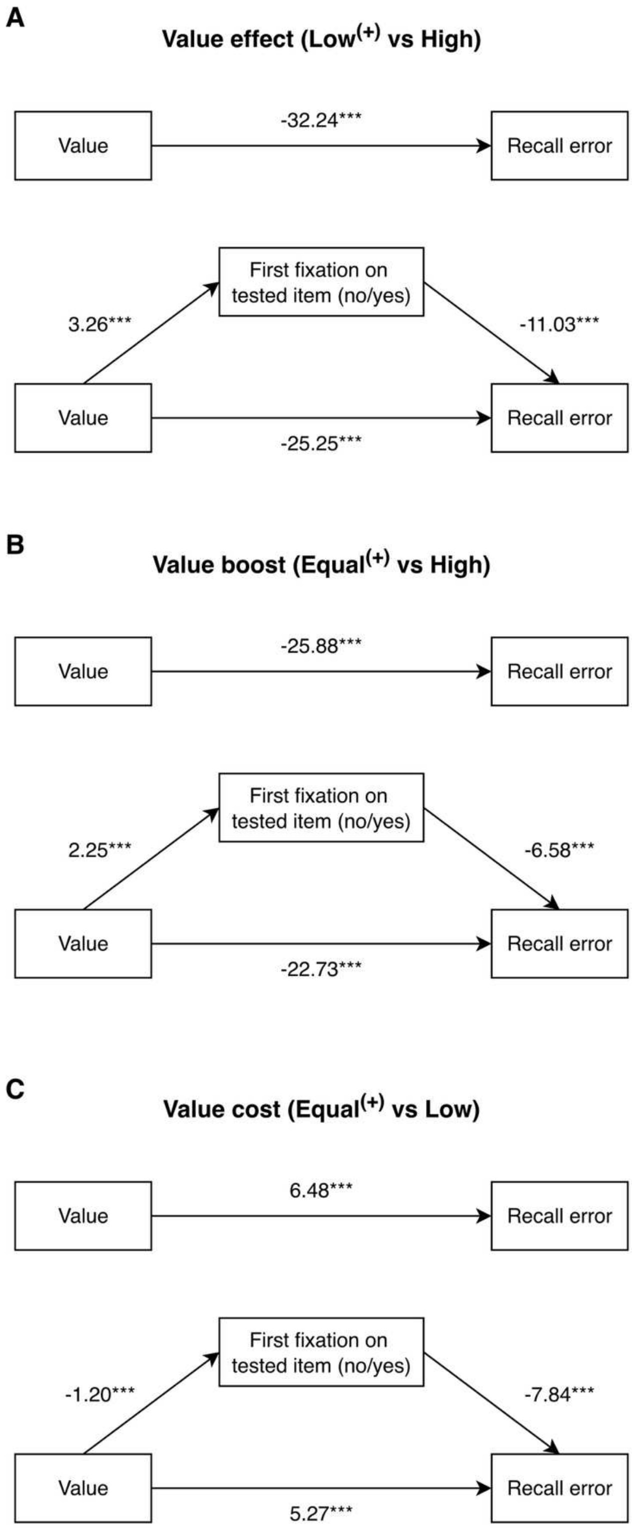
A mediation analysis investigated whether the first fixation participants made during encoding could explain the value-based prioritization effect. We examined the binary factor of whether participants had made their first fixation to the tested item (no/yes) as a mediator for the effect of value on recall error. As in the above analyses, one model was conducted for each value comparison (Low vs High, Equal vs High, Equal vs Low). The outcomes are presented in Figure 7. In all three models[3], the effect of value on recall error was partially mediated by whether the first fixation was directed towards the tested item. The indirect effects were -6.99 for the overall value effect (Low vs High; 95% CI: -9.29 - -4.70; $p < .001$), -3.15 for the value boost (Equal vs High; 95% CI: -5.03 - -1.35; $p = .001$), and 1.22 for the value cost (Equal vs Low; 95% CI: 0.60-1.90; $p < .001$). The proportion of the effect mediated was 0.22 (95% CI: 0.14-0.29; $p < .001$) for the overall value effect (Low vs High), 0.12 (95% CI: 0.05-0.20; $p = .001$) for the value boost (Equal vs High), and 0.19 (95% CI: 0.09-0.39; $p < .001$) for the value cost (Equal vs Low). Taken together, these outcomes indicate the effect of value on recall error was partially mediated by the first fixation participants made during encoding.[4]

Figure 7

Outcomes from the Mediation Analysis in Experiment 1, Investigating First Item fixated as the Mediator

[3] For the Low vs High model, the regression predicting first fixation from value (i.e. predictor -> mediator) was singular, due to 0 variance for the random intercept of ID. As it is not appropriate to compute some paths as mixed-effect models and other as standard regressions, we proceeded to run and present the model including ID. The same conclusions were drawn when excluding the random intercept of ID from all models for this comparison (Low vs High).

[4] We also examined whether encoding fixation duration still partially mediates the relationship between value and recall error when trials were excluded if participants fixated on the tested item first in the encoding phase. This enabled us to establish whether total encoding fixation duration still mediated the effect of value, after taking into account a primacy account. Partial mediation was still observed in all models (see supplementary materials for full reporting).



Note. Panel A reflects the overall value effect (Low vs High value), Panel B reflects the value boost (Equal vs High value), and Panel C reflects the value cost (Equal vs Low value). In all panels, the upper diagram reflects the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. In all panels, paths to "first fixation on tested item (no/yes)" are from logistic regression (with coefficients presented on the log odd scale). All other coefficients are from linear regression coefficients. $^{+}$ denotes reference category. *** $p < .001$, ** $p < .01$, * $p < .05$.

Discussion

A clear value effect was observed on recall error, whereby participants exhibited lowest recall error for high value items, intermediate recall error for equal value items, and the highest recall error for low value items. This is in line with previous studies, which have observed value-based prioritization effects in working memory, using continuous response tasks (Atkinson et al., 2022; Hu et al., 2023) and other retrieval tasks (e.g. cued-recall, recognition; Atkinson et al., 2018; 2021; 2022). The mixture modelling also revealed a number of useful insights. For example, participants had a higher probability of recalling high value items relative to other items, and were also more precise in their responses to high value items. These outcomes are broadly in line with the findings reported by Atkinson et al. (2022).

To the best of our knowledge, this was the first study to utilize eye-tracking to examine the mechanisms underlying value-based prioritization effects in working memory. This revealed several novel findings. Firstly, as predicted, the amount of time participants spent fixating on an item during encoding differed as a function of value. Participants fixated the longest on high value items, an intermediate amount of time on equal values items, and the shortest amount of time on low value items. Further, mediation analysis demonstrated that the amount of time participants spent fixating on the tested item during encoding partially explained the effect of value on recall error, accounting for approximately 36% of the overall value effect (low value vs high value), 29% of the value boost (equal value vs high value) and 36% of the value cost (equal value vs low value). Previous research has suggested that eye movements towards an item results in a greater proportion of memory resource being assigned to it (Bays & Husain, 2008). Thus, biases in eye movements towards high value information likely result in these items being encoded more strongly into working memory. In the current study participants also tended to fixate on the high value item first, with this factor also partially mediating the effect of value. The value-based prioritization effect therefore partially appears to reflect a primacy-like effect[5], whereby participants show a tendency to encode the high value item first. Taken together, these findings provide the first evidence that participants

[5] We term this a "primacy-like effect" to avoid confusion given that we presented information simultaneously and the term "primacy" is usually used in tasks involving sequential presentation.

vary their eye movement during encoding based on value, and that such behaviors partially explain the value effects robustly observed in working memory tasks (Allen et al., 2025).

Although most participants exhibited value-based prioritization effects, there was considerable variability in the size of the effects observed, with some participants showing a markedly larger effect than others. This is likely to reflect the highly strategic nature of value-based prioritization (Hu et al., 2016), with the size of effects influenced by numerous factors including the extent to which participants engage with the points system and prioritization instructions (Allen et al., 2025). A similar, variable pattern, was also observed in the eye movement analyses, though the vast majority of participants showed a bias towards the high value items to at least some extent.

Experiment 2

Experiment 1 indicated that participants show a bias towards high value information during the encoding phase, with total fixation time and first fixation both partially mediating the effects of value on task performance. As described earlier, there is evidence using cued-recall tasks that individuals can direct their attention in a highly sophisticated manner when all items differ in their value (e.g. items vary in their point values between 1-4; Allen & Ueno, 2018). Experiment 2 examined whether participants also vary their fixations during the encoding phase during these highly graded prioritization tasks. Participants completed a visual working memory task, similar to that employed in Experiment 1. On each trial, all items differed in their notional value (varying between 1-4 points). Of particular interest was how eye movements during encoding differ as a function of value, and whether the value effect is driven by eye movements during encoding. The manipulation of value in such a way also enables further, more in-depth analysis of the eye-tracking outcomes than was possible in Experiment 1. That is, varying the value of each item within a trial allows exploration of whether the order in which participants sampled the items was affected by value. In Experiment 1, we examined whether the proportion of time participants fixated on the high value item first was greater than would be expected based on an equal distribution. However, more complex analysis comparing value levels across the trial would have been confounded in Experiment 1, as there were three times as many low value items as high value items on each differential value trial. Thus, at any given position (first item fixated, second item fixated, etc.), participants would have been more likely to fixate on an item of low value, simply because there were more low value items available for sampling per trial relative to high value items. In contrast, assigning a different value to each item within a trial means that there would be no inherent bias in sampling towards items of a particular condition. This analysis provides novel insights into how patterns of fixations change over the encoding phase. More generally, use of a continuous-reproduction task (as in Experiment 1) also allows exploration of how graded value affects recall error and mixture modelling parameters, providing further novel insights into the value-based prioritization effect.

Based on previous research (e.g. Allen & Ueno, 2018; Allen et al., 2025), it was predicted that participants would show a value-based prioritization effect, with lower recall error for items that were associated with higher values. Several other predictions were made based on the outcomes of Experiment 1. It was predicted that the extent to which participants fixated on an item during encoding would differ as a function of value, with participants fixating on higher value information for a greater proportion of the encoding phase relative to lower value information. It was also predicted that the proportion of time participants spent fixating on an item during encoding would partially explain value-based prioritization effects. Moreover, based on the outcomes of Experiment 1, it was also anticipated that participants would often fixate on the highest value information first, and that first fixation would partially mediate the effect of value. Further, novel analyses then examined whether the following outcomes were affected by item value: the position each item was first sampled, the order items were sampled, and the proportion of trials in which participants re-sampled items.

Method

Participants

A power calculation was conducted in G*Power (Faul et al., 2007) for a repeated measures one-way ANOVA with four value levels (point values 1-4). Based on a medium effect size (F = 0.25) and alpha set to 0.05, a sample of 30 participants was required to provide 90% statistical power. A new sample of thirty-nine participants completed the experiment. All participants met the same criteria as outlined in Experiment 1. All participants provided demographic information (*M.age* = 20.97, *SD* = 2.92; Gender: 25 women, 13 men, 1 non-binary). One participant was excluded, as the percentage of trials with no usable eye-tracking data was 2.5 standard deviations above the mean. The final analysis was therefore performed on data from 38 participants (*M.age* = 20.97, *SD* = 2.96; Gender: 24 women, 13 men, 1 non-binary).

Design, Materials, and Procedure

The experimental paradigm is displayed in Figure 1B. The design, materials, and procedure were identical to Experiment 1, except for the following adaptations. Value remained the independent variable, but with four levels (1-4 points) in this experiment. On every trial, there was one item worth 1 point, one worth 2 points, one worth 3 points, and one worth 4 points (as in Allen & Ueno, 2018). Further, as the array of point values was more complex in this experiment, the display time for the point value screen was increased to 3500ms. All other timings were the same as in Experiment 1, except that the retention interval was very marginally longer across participants due to random variation (*M* across participants = 2132ms (vs 2130ms in Experiment 1), *SD* across participants = 7ms; Range across participants: 2118-2152ms). Participants completed 192 test trials in total. Each of the four value items (1-2-3-4 points) were tested on 48 trials, with trial order intermixed. Within each value level, each of the four spatial locations were tested equally. Participants completed four practice trials at the start of the experiment which were not analyzed.

Data Analysis

Analyses were performed as outlined in Experiment 1. In the Bayesian three-component mixture model, there was evidence of good convergence (all $\hat{R}$ = 1.00).

In the eye-tracking data, the mean percentage of missing samples during the encoding phase in the final sample (*N* = 38) was 1.65% (*SD* = 1.91). Trials where no fixations were recorded during the encoding phase were excluded (M = 0.81%, *SD* = 1.53). Following this, the mean number of fixations across participants in the encoding phase was 4.73 (*SD* = 0.81; *SE* = 0.13, Range = 2.89-5.96).

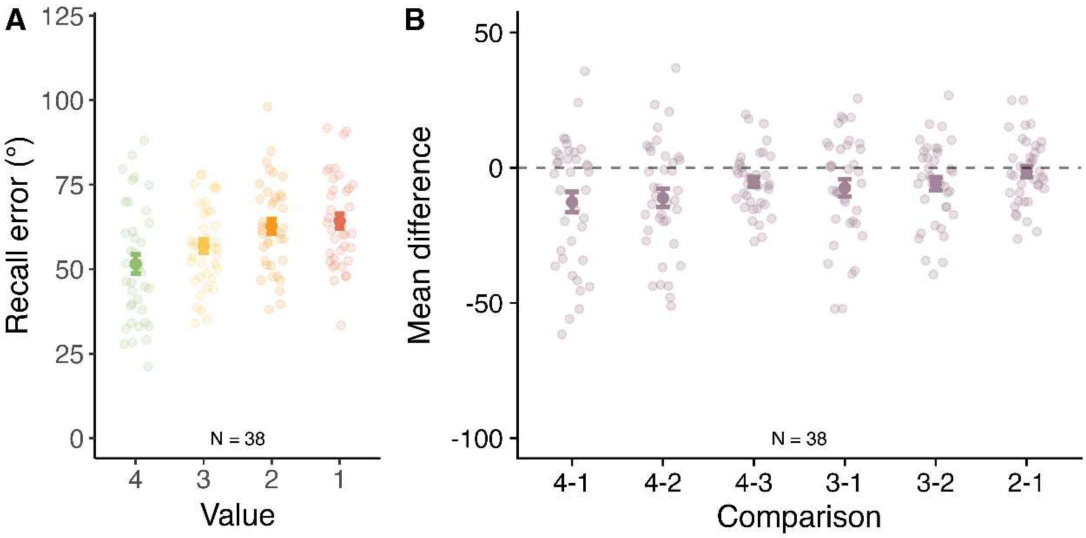
Results

Recall Error

Mean recall error (and SE) is displayed in Figure 8A as a function of value. A repeated measures one-way ANOVA on the factor of value (four levels: 1-4 points) was significant (*GG-corrected F*(1.67, 61.68) = 8.23, *MSE* = 281.13, $p < .001$, $\eta_p^2 = .18$, $BF_{10} = 558.39$). Post-hoc tests revealed that recall error for items worth 4 points ($M = 51.5$, $SE = 2.84$) was significantly lower than items worth 1 point ($M = 64.2$, $SE = 2.19$, $p = .012$, $d = -0.54$; $BF_{10} = 17.34$), 2 points ($M = 62.7$, $SE = 2.17$; $p = .012$, $d = -0.53$; $BF_{10} = 15.29$) and 3 points ($M = 56.8$, $SE = 1.99$; $p = .024$, $d = -0.47$; $BF_{10} = 6.37$). No other significant differences were observed ($ps \geq .062$, $ds \leq -0.12$ - -0.39). BF analysis provided weak evidence of a difference between items worth 3 points and 1 point ($BF_{10} = 1.99$), and 3 points and 2 points ($BF_{10} = 2.24$). There was BF evidence against a difference between items worth 1 and 2 points ($BF_{10} = 0.23$).

Figure 8

Task performance in Experiment 2



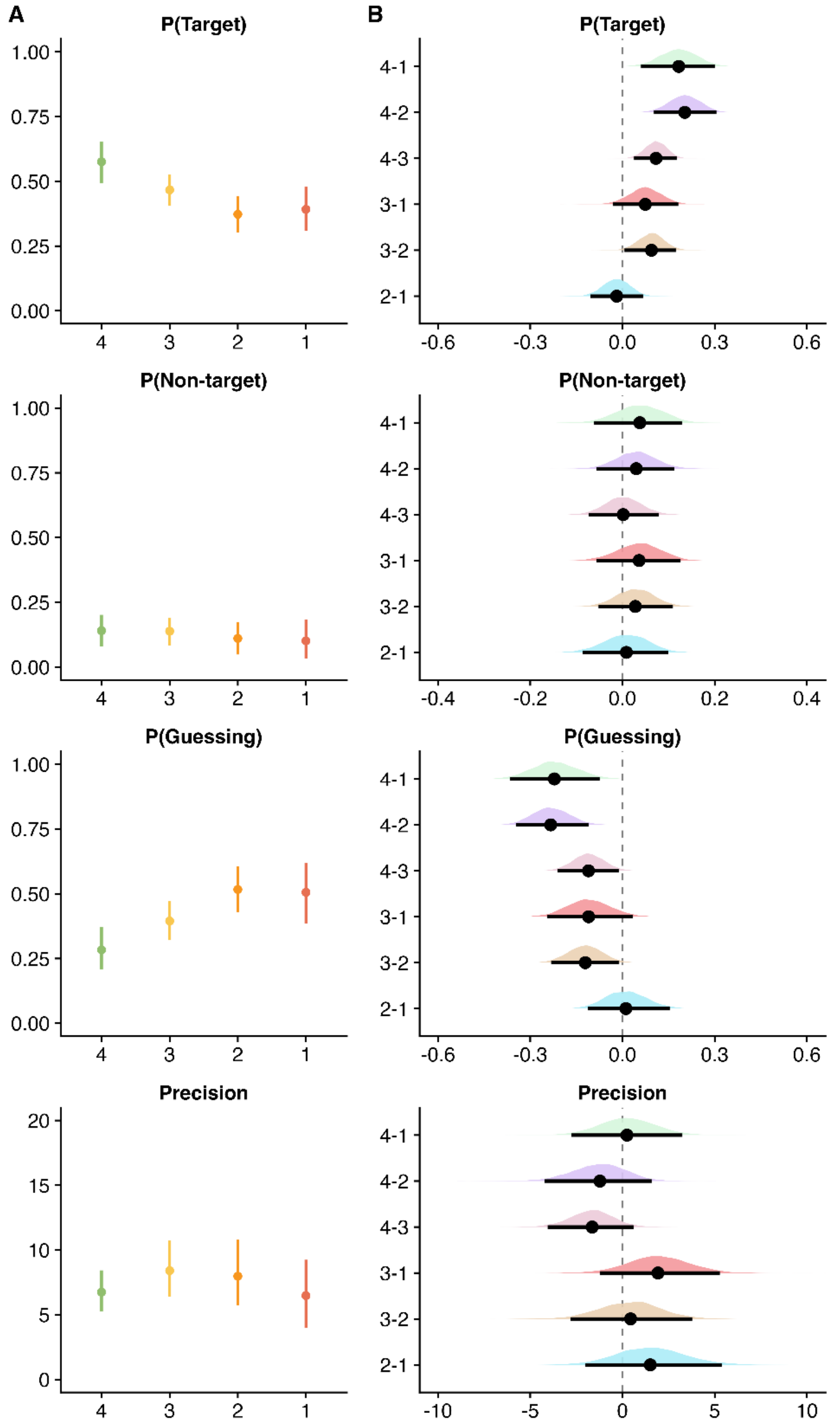
Note. Panel A displays mean recall error, as a function of value (higher recall error reflects worse performance). Panel B displays the mean difference between the value conditions. The dotted line at 0 in Panel B reflects no difference. Values above 0 indicate higher recall error (i.e. worse performance) in the first condition listed, whilst values below 0 indicate lower recall error (i.e. better performance) in the first condition listed. Across panels, bold points reflect mean across participants, whilst error bars denote standard error. Lighter circles in the background indicate individual participants' mean.

Bayesian Mixture Modelling

The parameter estimates from the Bayesian mixture model are displayed in Figure 9A, as a function of value. The differences in posterior distributions were calculated between all value conditions. The median, 95% highest density intervals, and distributions are presented in Figure 9B. The modelling revealed that the probability of recalling the target item was higher for items associated with 4 points relative to those worth 1 point (*pd* = .998), 2 points (*pd* = 1.000) and 3 points (*pd* = .999). Moreover, the probability of recalling the target item was higher for items associated with 3 points relative to those associated with 2 points (*pd* = .983). There were no credible differences between the other conditions in the probability of recalling the target item (*pd* ≤ .914). The probability of guessing was lower for items associated with 4 points relative to those associated with 1 point (*pd* = .998), 2 points (*pd* = 1.000), and 3 points (*pd* = .983). The probability of guessing was also lower for items associated with 3 points relative to 2 points (*pd* = .983). No other credible differences were observed regarding the probability of guessing (*pd* ≤ .935). Finally, there were no credible differences between value conditions when considering the probability of recalling a non-target item (*pd* ≤ .775) or precision (*pd* ≤ .929).

Figure 9

Outcomes from the Bayesian Mixture Modelling in Experiment 2



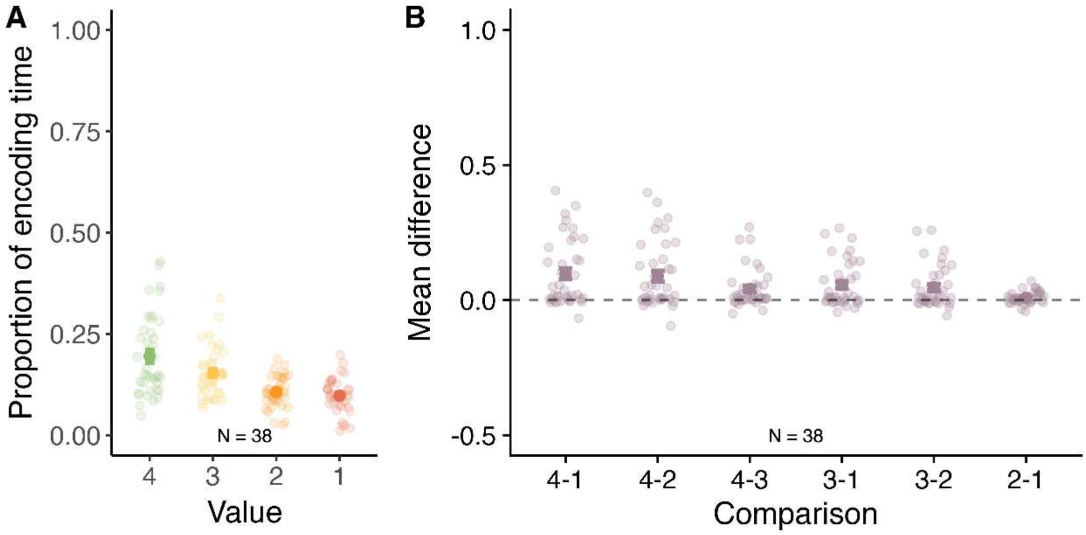
Note. Panel A: Estimates from the Bayesian mixture modelling as a function of value. Points reflect the median, whilst error bars display 95% highest density intervals. Panel B: Difference between the posterior distributions for the parameters in the Bayesian mixture modelling. For each parameter, differences are displayed for all comparisons between the value conditions. The black circle displays the median, the black horizontal bar shows the 95% highest density intervals, and the color shading reflects the posterior distribution. The grey vertical line at 0 reflects no difference. In both panels, facets from top to bottom show: probability of recalling the target item, probability of recalling a non-target item, probability of guessing, and precision.

Eye-Tracking Outcomes

Encoding Fixation Duration. The proportion of time participants spent fixating on each item during the encoding period is presented in Figure 10A, as a function of item value. Figure 10B displays the differences in the proportion of encoding time between different item value conditions. A one-way repeated measures ANOVA on the factor of value (four levels: 1-4 points) was significant (GG-corrected $F(1.32, 48.98) = 19.11$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .34$, $BF_{10} > 10{,}000$). Post-hoc tests indicated that participants spent a significantly higher proportion of the encoding time fixating items worth 4 points ($M = .20$, $SE = .02$), relative to those worth 1 point ($M = .10$, $SE = .01$; $p < .001$; $d = 0.79$; $BF_{10} = 1033.98$), 2 points ($M = .11$, $SE = .01$; $p < .001$; $d = 0.73$; $BF_{10} = 351.42$), and 3 points ($M = .15$, $SE = .01$, $p = .002$, $d = 0.57$; $BF_{10} = 28.70$). Participants also spent a greater proportion of time fixating items worth 3 points relative to those worth 1 point ($p = .001$, $d = 0.67$; $BF_{10} = 129.19$) or 2 points ($p = .002$, $d = 0.61$; $BF_{10} = 47.78$). Finally, participants spent a greater proportion of time fixating items worth 2 points relative to those worth 1 point ($p = .016$, $d = 0.41$), though BF support was weak ($BF_{10} = 2.78$).[6]

Figure 10

The Proportion of the Encoding Phase Participants Spent Fixating on Each Item in Experiment 2, as a Function of Value



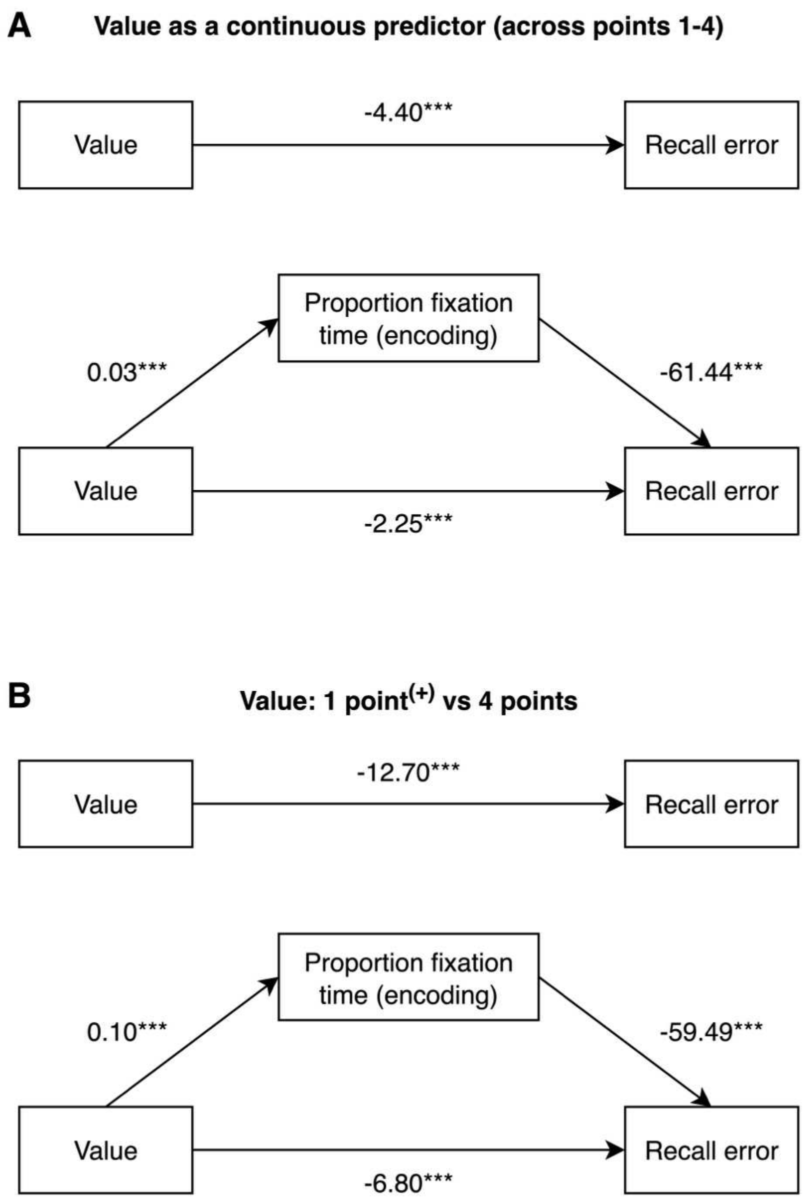
[6] Further analysis was conducted to establish whether this effect was driven by relatively more fixations towards high value items or individual fixations being longer towards high value items. Both of these patterns were observed, suggesting participants fixate on the high value item more often and longer relative to lower value items. See supplementary materials for full reporting.

Note. Panel A displays the mean proportion of the encoding time participants spent fixating on an item, as a function of value. Panel B displays the mean difference between the value conditions. The dotted line at 0 in Panel B reflects no difference. Values above 0 indicate that participants fixated on the first condition for a greater proportion of the encoding phase relative to the second level listed. Meanwhile, values below 0 indicate that participants fixated on the first condition listed for a smaller proportion of the encoding phase relative to the second level listed. Across panels, bold points reflect the mean across participants, whilst error bars denote standard error. Lighter circles in the background indicate individual participants' means.

Next, mediation analysis was conducted to investigate whether the proportion of time participants fixated on the tested item during encoding mediated the relationship between value and recall error. As in Experiment 1, this analysis included only trials that had any fixations detected during the encoding phase. Given that there were various value conditions in this experiment that increased numerically (i.e. from 1-4), value was entered as a continuous predictor into this model. As shown in Figure 11A, all relationships were significant. The indirect effect was -2.15 (95% CI: -2.50 – -1.81; $p < .001$), and the proportion of the effect mediated was 0.49 (95% CI: 0.38-0.66; $p < .001$). A second mediation model was then conducted, considering only point values "1" and "4" (to be broadly comparable to the "Low" and "High" value analyses conducted in Experiment 1). The outcomes are presented in Figure 11B. All relationships were significant in this model. In this model, the indirect effect was -5.90 (95% CI: -7.20 – -4.72; $p < .001$), and the proportion of the effect mediated was 0.46 (95% CI: 0.34-0.67, $p < .001$). Thus, in both models, the effect of value on recall error was partially mediated by the proportion of time participants fixated on the tested item during encoding.

Figure 11

Outcomes from the Mediation Analysis in Experiment 2, Examining Proportion of Fixation Time during Encoding as the Mediator

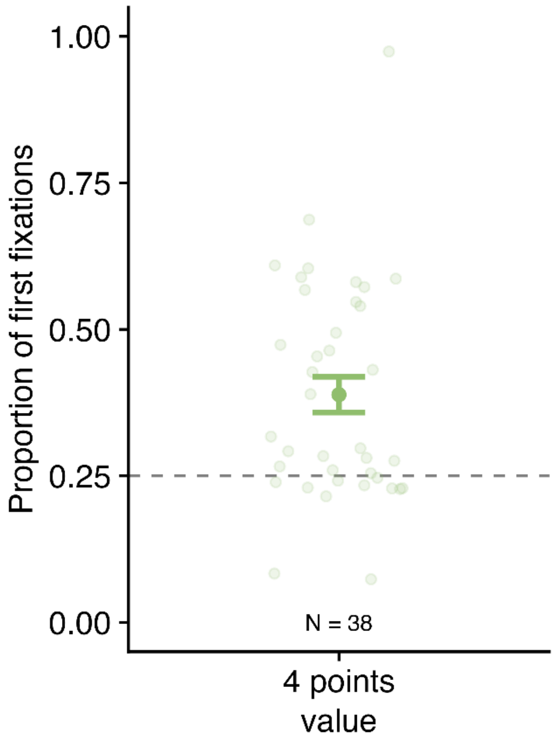


Note. Panel A: Outcomes from the mediation model, when considering all four value conditions (values 1-4), entered into the model as a continuous predictor. Panel B: Outcomes from a mediation model, comparing "1" and "4" point values (comparable to the "Low" vs "High" mediation analyses in Experiment 1). In this latter model, the "1" point condition was entered as the reference category (denoted in the figure by (+)). In both panels, the upper diagram displays the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. *** $p < .001$, ** $p < .01$, * $p < .05$.

First Item Fixated. As in Experiment 1, we investigated whether the proportion of trials in which participants fixated on the highest value item (i.e. that associated with 4 points) first exceeded that expected based on an equal distribution of fixations (0.25). Trials were excluded where participants did not fixate on any of the items during the encoding phase. The mean proportion of trials in which participants fixated first on the item worth 4 points is displayed in Figure 12. A one sample t-test indicated that the mean proportion of fixations towards the item worth 4 points items (M = 0.39, SE = .03) was significantly greater than expected based on an equal distribution (0.25), $t(37) = 4.55$, $p < .001$, $d = 0.74$, $BF_{10} =$ 411.13).

Figure 12

Mean Proportion of Trials in which the First Item Fixated was the Highest Value Item in Experiment 2



Note. The dotted line at 0.25 reflects an equal distribution of first fixations across items. The bold point reflects the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error of the mean.

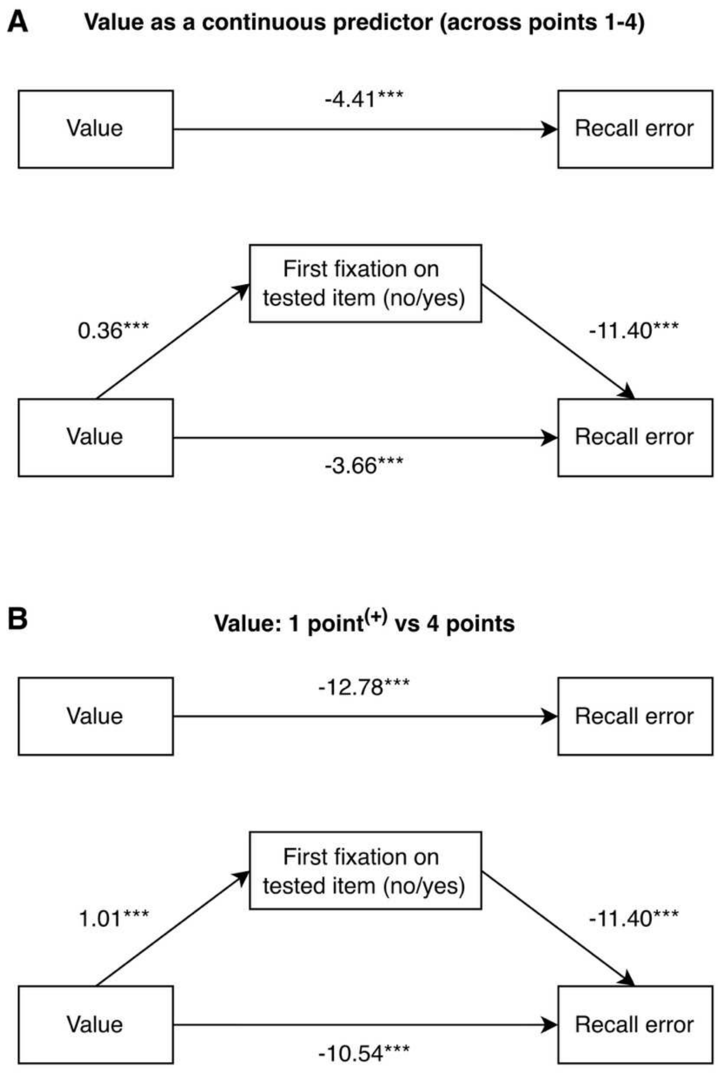
A pair of mediation models were then conducted to investigate whether the first item participants fixated on (towards tested item – no/yes) mediated the relationship between value and recall error. As in the main analyses, the first model investigated value as a continuous variable (1-4), whilst the second compared values 1 and 4. The outcomes are presented in Figure 13 for both models[7]. For the model with value as a continuous variable,

[7] For the model with value as a continuous variable, the regression predicting first fixation from value (i.e. predictor -> mediator) was singular, due to variance for the random intercept of ID being very close to 0. As it is

the indirect effect was -0.47 (95% CI: -0.66 - -0.31; $p < .001$), and the proportion mediated was 0.11 (95% CI: 0.07-0.18; $p < .001$). For the model comparing values 1 vs 4 only, the indirect effect was -2.25 (95% CI: -3.20 - -1.42; $p < .001$), and the proportion mediated was 0.17 (95% CI: 0.11-0.28; $p < .001$).[8]

Figure 13

Outcomes from the Mediation Analysis in Experiment 2, Investigating First Fixation as the Mediator



not appropriate to compute some paths as mixed-effect models and other as standard regressions, we proceeded to run and present the model including ID. The same conclusions were drawn when excluding the random intercept of ID from all models.

[8] As in Experiment 1, we also examined whether encoding fixation duration still partially mediated the relationship between value and recall error when trials were excluded if participants fixated on the tested item first in the encoding phase. Partial mediation was observed in both models (see supplementary materials for full reporting).

Note. Panel A: Outcomes from the mediation model, when considering all four value conditions (values 1-4), entered into the model as a continuous predictor. Panel B: Outcomes from a mediation model, comparing "1" and "4" point values (comparable to the "Low" vs "High" mediation analyses in Experiment 1). In this latter model, the "1" point condition was entered as the reference category (denoted in the figure by (+)). In both panels, the upper diagram displays the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. *** $p < .001$, ** $p < .01$, * $p < .05$.

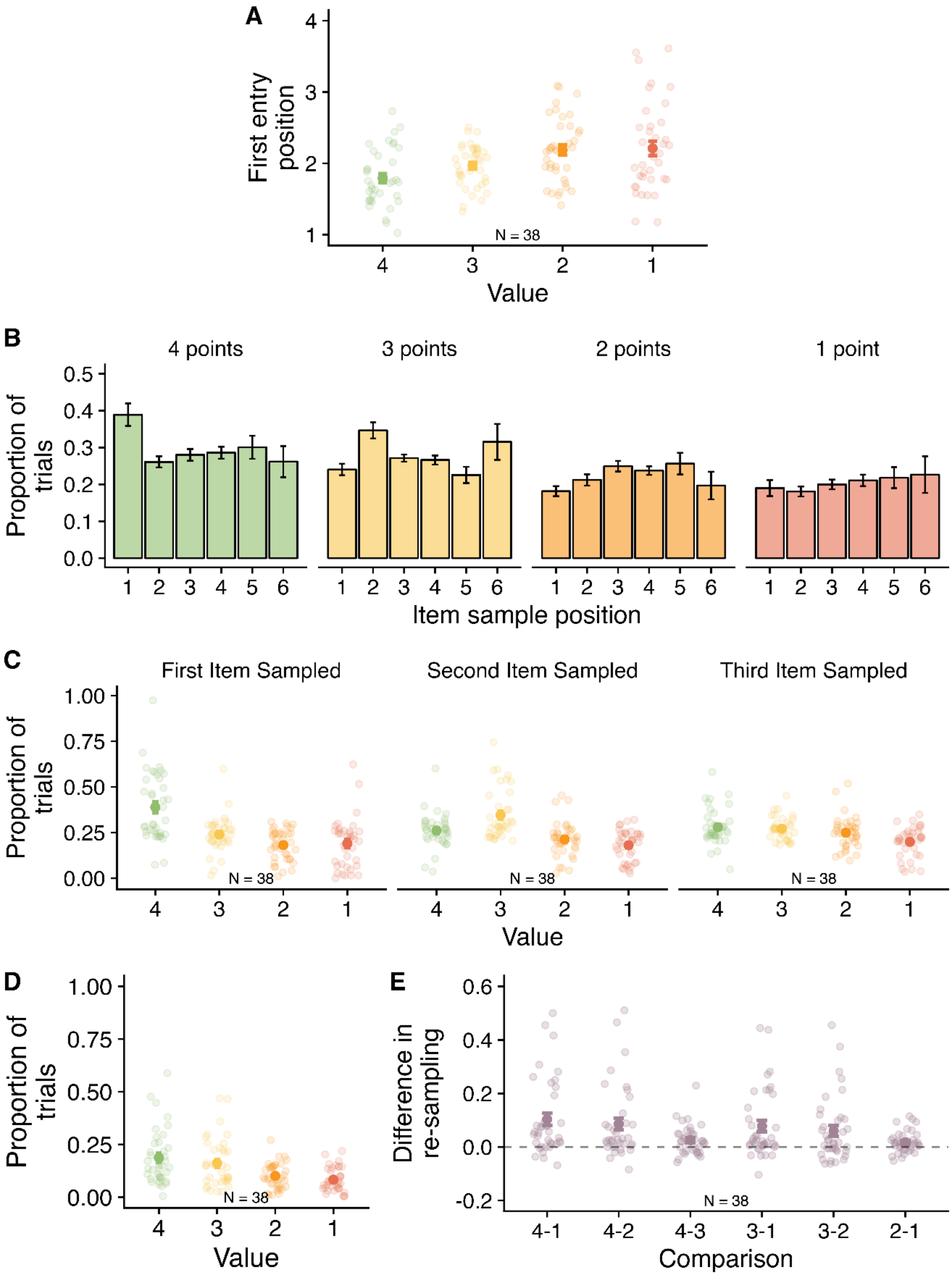
Subsidiary Analyses.

Position of First Sample by Item Value. This analysis examined when each item was first sampled in the sequence of entries across the encoding phase, and whether this differed as a function of item value. A value of 1 on this metric would indicate participants always accessed that item value first on every trial. This figure ignores trials in which an entry was not made to an item of a particular value. This provides a measure of sampling position that is unconfounded by differences in sampling rates across value levels. Figure 14A shows the mean position in which an item was first sampled, as a function of item value.

A one-way ANOVA revealed a significant effect of value (GG-corrected $F(1.29, 47.68) = 9.26$, $MSE = 0.37$, $p = .002$, $\eta_p^2 = .20$, $BF_{10} = 1817.42$). Post-hoc tests revealed that the mean position in which an item was first sampled was lower for items worth 4 points ($M = 1.79$, $SE = 0.06$) than those worth 1 point ($M = 2.21$, $SE = 0.10$; $p = .015$, $d = -0.80$, $BF_{10} = 9.54$), 2 points ($M = 2.19$, $SE = 0.07$; $p = .003$, $d = -0.93$, $BF_{10} = 68.82$), and 3 points ($M = 1.97$, $SE = 0.05$; $p = .007$, $d = -0.50$, $BF_{10} = 22.55$). Moreover, the mean position of first sampling was lower for items worth 3 points than items worth 2 points ($p = .015$, $d = -0.58$, $BF_{10} = 7.60$). No significant difference were observed between items worth 2 points and 1 point ($p = .675$, $d = -0.04$, $BF_{10} = 0.19$). There was also no significant difference between items worth 3 points and 1 point ($p = .055$, $d = -0.50$), though there was weak BF support ($BF_{10} = 1.77$).

Figure 14

Outcomes of the Subsidiary Analyses in Experiment 2



Note. Panel A: Position of first sample by item value. Panel B: The proportion of trials in which participants fixated on items, as a function of value and item sample position (first item attended; second item attended, etc). Data is presented for the first six item sample positions only, as few participants had data beyond this. Panel C: The proportion of trials in which participants fixated on items as a function of value for the first three item sample positions only. Panel D: The mean proportion of trials involving re-sampling of items, as a function of item value. Panel E: The mean difference in re-sampling between value pairs. The *dotted horizontal line at 0 reflects no difference. Values above 0 indicate that participants re-sampled items in the first condition to a greater extent relative to the second. Meanwhile, values below 0 indicate that participants' re-sampled items in the first condition to a lesser extent than the second. In all, error bars denote standard error. In panels A, C, D and E, lighter dots in the background reflect means for individual participants.*

Order of Item Sampling. We next examined whether the order in which items were sampled differed as a function of item value. Trials were first excluded where there were no fixations to any of the items. For each trial, we then determined the sequence of area of interest (AOI) entries, and then for each item value (1-4) computed the position in which it was sampled. We term this "Item sample position". For each participant, the proportion of trials in which items were sampled was computed for each value and item sample position combination. Here, the denominator was the number of trials in which participants sampled any item at the given sample position. Data was then averaged across participants. The number of items sampled varied considerably across trials and participants. Not all participants contributed data for all item sample positions. The mean proportion of trials in which items were sampled is displayed in Figure 14B, as a function of item value and item sample position. Data is presented for the first six item sample positions only, as few participants had data beyond this. Of note, this figure shows that items associated with 4 points were most likely to be fixated first (in almost 40% of trials), while items associated with 3 points were most likely to be fixated second in the sequence.

Data was analyzed at the first three item sample positions, as the vast majority of participants (37/38) sampled all item values to some degree across these positions. Mean proportion of fixations at each of the first three item sample positions for this sample is presented in Figure 14C, as a function of item value. A 4 (Value: 1-4) x 3 (Item sample position: first, second, or third) within-subjects ANOVA revealed a significant effect of item value (GG-corrected $F(1.27, 46.85) = 19.34$, $MSE = 0.05$, $p < .001$, $\eta_p^2 = .34$, $BF_{10} > 10{,}000$). Post-hoc tests revealed significant differences between all comparisons (4 points $M = .31$, $SE = .01$; 3 points $M = .29$, $SE = .01$; 2 points $M = .21$, $SE = .01$; 1 points $M = .19$, $SE = .01$; $p \leq .011$, $d \geq 0.19$). There was also very strong BF support for all comparisons ($BF_{10} \geq 1322.29$), except for the comparison between 4 points and 3 points where $BF_{10} = 0.24$, and the comparison between 2 points and 1 point where $BF_{10} = 1.87$. There was no significant effect of item sample position ($F(2, 74) = 0.00$, $MSE < .01$, $p = 1.000$, $\eta_p^2 < .01$, $BF_{10} = 0.02$). Crucially, there was a significant interaction between value and item sample position (GG-corrected $F(2.49, 92.04) = 8.23$, $MSE = 0.03$, $p < .001$, $\eta_p^2 = .18$, $BF_{10} > 10{,}000$).

To break down the interaction, a one-way ANOVA was conducted at each item sample position. At item sample position 1 (i.e. first fixated item), there was a significant effect of item value (GG-corrected $F(1.62, 59.87) = 14.97$, $MSE = 0.04$, $p < .001$, $\eta_p^2 = .29$, $BF_{10} > 10{,}000$). Post-hoc tests revealed that participants fixated on the item worth 4 points ($M = .39$, $SE = .03$) first on a greater number of trials than items worth 1 point ($M = .19$, $SE = .02$; $p = .001$; $d = 1.21$, $BF_{10} = 92.02$), 2 points ($M = .18$, $SE = .01$; $p < .001$; $d = 1.42$; $BF_{10} = 1108.61$), or 3 points ($M = .24$, $SE = .02$; $p = .001$; $d = 0.99$, $BF_{10} = 127.28$). Participants also fixated on items worth 3 points first on a higher proportion of trials than items worth 2 points ($p = .039$; $d = 0.65$, $BF_{10} = 3.30$). No significant differences emerged between items worth 3 points and 1 point ($p = .245$; $d = 0.43$, $BF_{10} = 0.54$), or items worth 2 points and 1 point ($p = .626$; $d = -0.07$, $BF_{10} = 0.20$).

At item sample position 2 (i.e. second item fixated), a significant effect of value also emerged (GG-corrected $F(2.10, 77.65) = 13.96$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .27$, $BF_{10} > 10{,}000$). Bonferroni-Holm corrected post-hoc tests revealed that participants fixated on the item worth 3 points ($M = .35$, $SE = .02$) second on a greater proportion of trials than items worth 1 point ($M = .18$, $SE = .01$; $p < .001$; $d = 1.48$; $BF_{10} = 1422.16$), 2 points ($M = .21$, $SE = .02$; $p = .001$; $d = 1.16$; $BF_{10} = 101.65$), or 4 points ($M = .26$, $SE = .02$; $p = .014$; $d = 0.74$, $BF_{10} = 8.15$). Participants also fixated on items worth 4 points second on a greater proportion of trials than items worth 1 point ($p = .004$; $d = 0.91$; $BF_{10} = 29.10$). No significant differences emerged between items worth 4 points and 2 points ($p = .110$; $d = 0.52$; $BF_{10} = 0.80$), or items worth 2 points and 1 point ($p = .110$; $d = 0.35$; $BF_{10} = 1.01$).

Finally, at item sample position 3 (i.e. third item fixated), there was significant effect of value (GG-corrected $F(1.93, 71.35) = 5.34$, $MSE = 0.01$, $p = .008$, $\eta_p^2 = .13$, $BF_{10} = 142.90$). Bonferroni-Holm corrected post-hoc tests revealed that participants fixated on the item worth 1 point ($M = .20$, $SE = .01$) on a lower proportion of trials than items worth 4 points ($M = .28$, $SE = .02$; $p = .021$; $d = -0.89$, $BF_{10} = 8.70$), 3 points ($M = .27$, $SE = .01$; $p = .008$; $d = -1.00$, $BF_{10} = 23.64$), or 2 points ($M = .25$, $SE = .01$; $p = .024$; $d = -0.59$; $BF_{10} = 6.50$). No other significant differences emerged (p for all other comparisons = .873; $d \leq 0.33$, $BF_{10} \leq 0.30$).

Thus, the items participants first fixated were most likely to be the item worth 4 points, whilst the second item they fixated on was most likely to be the item worth 3 points. The outcomes at the third fixation were less clear, though there was evidence participants were less likely to fixate on items worth 1 point relative to the rest.

Re-sampling of Items. Finally, analysis was conducted examining whether the proportion of trials in which an item received two separate periods of inspection differed as a function of item value. Trials were excluded from this analysis where there were no fixations to any of the items. As this analysis was interested in *separate* periods of inspection, trials were not counted for a given item value if they only contained consecutive re-entries. For example, if participants sampled item value 4, then 3, then 4, the trial would be classified as a re-sampling occurrence for item value 4. Whereas, if participants sampled item values 4, then 4, then 3, the trial would not be counted as re-sampling of item value 4 (or indeed any item value condition). The mean proportion of trials where re-sampling occurred is displayed in Figure 14D, as a function of item value. Differences between value condition pairs are displayed in Figure 14E. A one-way ANOVA was conducted to investigate whether the proportion of trials with re-sampling differed as a function of item value. There was a significant effect of value (GG-corrected $F(1.26, 46.71) = 14.84$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = .29$, $BF_{10} > 10{,}000$), with post-hoc tests revealed significant differences between all conditions (4 points $M = .19$, $SE = .02$; 3 points $M = .16$, $SE = .02$; 2 points $M = .10$, $SE = .01$; 1 points $M = .08$, $SE = .01$; $ps \leq .012$, $ds \geq 0.21$, $BFs_{10} \geq 4.65$). Thus, participants re-sampled higher value items to a greater extent than lower value items, with a graded effect present across the levels.

Discussion

There was a clear effect of value on recall error. Such findings are in line with Experiment 1 and previous research examining effects of value on recall error (Atkinson et al., 2022; Hu et al., 2023) and other outcome measures (see Allen et al., 2025). There was some numerical evidence that recall error decreased as point values increased, although support from the inferential analyses was not particularly strong. In the frequentist analyses, there were only significant differences between items worth the highest number of points (4 points) and the other items, although the BF analysis also provided some weak support for a difference between items worth 3 points and those worth 2 or 1 point. The mixture modelling revealed that participants had a higher probability of recalling the target item in the 4-point condition relative to all other values, and also had a higher probability of recalling the target item in the 3-point condition relative to the 2 points condition. Taken together, these findings provide some evidence of a graded value effect, although this was not as clear as in previous work which has utilized a similar manipulation in cued recall paradigms (Allen & Ueno, 2018). Nevertheless, there was still clear evidence that participants were able to prioritize high value items relative to lower value items. Interestingly, unlike in Experiment 1 and past research utilizing mixture modelling to examine effects of value (Atkinson et al., 2022), no effects on precision were observed for any value pair comparisons. This may suggest that the effect of value on precision is only observed under less complex reward conditions (i.e. when only one item is more valuable than the rest; Experiment 1 in the present study; Atkinson et al., 2022; Hu et al., 2023).

A clear value effect also emerged on the proportion of time participants spent fixating on an item during encoding, with participants fixating on higher value items more than lower value items. Post-hoc tests revealed a highly graded effect (i.e. 4 > 3 > 2 > 1 points). These findings replicate and extend Experiment 1, indicating that participants allocate their attention differently during the encoding phase when items differ in their relative values. Mediation analyses were then conducted to examine whether the proportion of time participants spent fixating on an item mediated the effect of value on recall error. The first model examined value as a continuous predictor, whereas the second compared the lowest (i.e. 1 point) and highest (i.e. 4 point) value conditions. Across both analyses (and consistent with Experiment 1), the proportion of time participants spent fixating on an item during the encoding phase partially mediated the effect of value on recall error. Further, as in Experiment 1, participants fixated on the item worth 4 points first more than would be expected based on an equal distribution. First fixation partially mediated the effect of value, providing further evidence that the effect of value on recall error partially reflects a primacy-like effect.

The design implemented in Experiment 2 also enabled further in-depth analyses of eye-tracking parameters. This revealed several interesting insights. For example, the first item participants fixated on during the encoding phase was most likely to be the item worth 4 points whilst the second item participants fixated on was most likely to be worth 3 points (i.e.

the next highest value item). This provides further support for a primacy-based account, suggesting that at least at a group level, participants engaged in a strategic prioritization of sampling in which they focused early eye movements during the encoding phase on the two highest valued items. This conclusion was further supported by other analyses, indicating that the average position in which items were first sampled tended to be earlier for higher value items. There was also clear evidence that participants re-sampled higher value items to a greater extent relative to lower value items. Taken together, these analyses provide further, consistent evidence that participants' eye movements during encoding are guided by strategic goals of item prioritization.

Finally, as in Experiment 1, there were large individual differences in both recall error and the eye-tracking outcomes. This further supports the notion that both the size of value-based prioritization effects in working memory, and the strategies underlying such effects, are subject to individual differences. However, as in Experiment 1, most participants showed a bias towards higher value items across all eye tracking measures investigated.

General Discussion

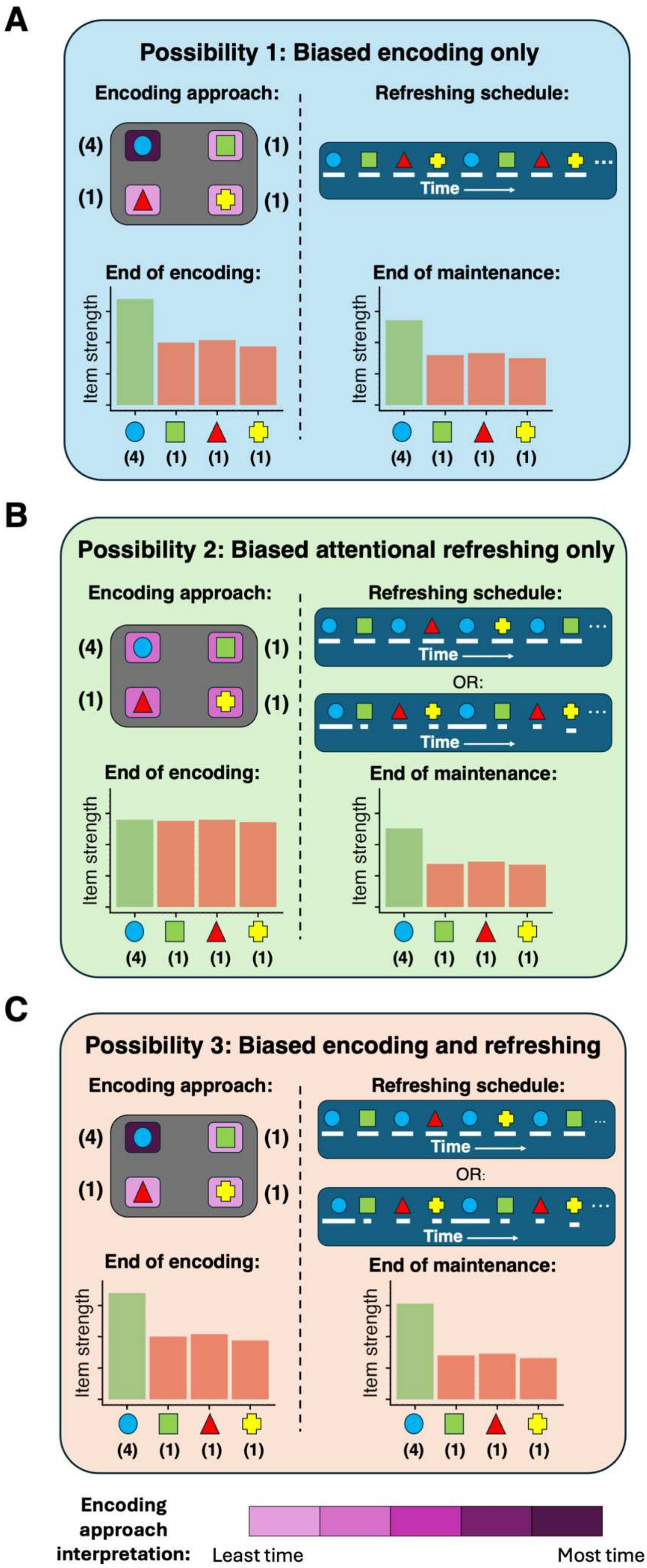
Studies examining mechanisms underlying value-based prioritization in working memory have focused on maintenance mechanisms, such as attentional refreshing and articulatory rehearsal (e.g. Atkinson et al., 2021; 2022; Sandry et al., 2014). Meanwhile, the role of encoding processes has received less attention and is consequently less well understood. Two experiments were conducted which investigated the extent to which: (i) patterns of eye movements during encoding vary as a function of item value; and (ii) whether eye movements during encoding drive value-based prioritization effects in working memory tasks. Experiment 1 used trials in which the to-be-remembered items were either of high or low value, and other trials in which all items were of equal value. Experiment 2 provided a more fine-grained analysis of the effect of value by varying the value of all items in the array. There was clear evidence of value-based prioritization effects, with participants exhibiting better performance for high value items relative to lower value items (e.g. Allen et al., 2025; Allen & Ueno, 2018). There was also clear evidence that participants' eye movements during encoding differed as a function of value, with such measures partially mediating the effect of value on recall error. Taken together, this pair of experiments provides novel evidence that: (i) participants alter their behaviors during the encoding phase of working memory tasks when items differ in their value, engaging in several behaviors that aim to preferentially encode the high value information; and (ii) eye movements during encoding partially mediate the effect of value on recall error.

How might eye movements during encoding drive value-based prioritization effects? Our study provides key insights into this. Firstly, participants often fixated on high value information first during encoding, with this factor partially mediating the effect of value on recall error. Thus, the value-based prioritization effect is partially explained by a primacy-like effect, whereby participants often fixate the highest value item first. The effect of value was also partially mediated by total fixation time during encoding, even after controlling for a primacy-based explanation. As encoding time increases, the strength with which an item is encoded is also considered to increase (Oberauer et al., 2012). Thus, fixating on high value items to a greater extent during encoding is likely to result in stronger representations of these items (Bays & Husain, 2008). Several factors are therefore likely to explain how eye movements during encoding drive value-based prioritization effects.

Previous research has suggested that biased attentional refreshing during maintenance at least partially mediates the effect of value (Atkinson et al., 2022; Sandry et al., 2014). The current study provides important clarification on this, demonstrating that biased attentional refreshing is unlikely to be the *sole* mechanism driving value-based prioritization effects. Thus, it is likely that encoding factors (e.g. primacy-like effects, increased encoding strength) and maintenance mechanisms (e.g. biased attentional refreshing) work in parallel to drive value-based prioritization effects. Participants encode high value items first and more strongly relative to other items, and are more likely to return to these items for additional encoding. These items are then refreshed for longer and/or more frequently during maintenance (Atkinson et al., 2022; see Figure 15). An interesting outstanding question is the extent to which processes at encoding and maintenance are separable. Participants may actively prioritize high value information during both encoding and maintenance as a direct result of value, and an attempt to maximize task performance. Other possibilities exist, however. For example, it is possible that attentional refreshing may be biased towards the high value item (Atkinson et al., 2022) purely because that item is represented most strongly at the end of encoding. In tasks where all items are equally valuable and goal-relevant, it has been suggested that attentional refreshing is targeted towards the item with the weakest level of activation (Camos et al., 2018; Lemaire et al., 2018). However, this may not be the case if the items that have the weakest level of activation are also of low value. In such circumstances, individuals may take a different approach; either abandoning such items all together (i.e. not refreshing them at all) or further de-prioritizing them by refreshing them less or for shorter periods of time. This would leave more time to refresh the higher value items in the sequence. A combination of these two possibilities would also be plausible, whereby biased attentional refreshing during maintenance reflects both the item being particularly valuable *and* being represented most strongly at the end of encoding. Further research should attempt to delineate between these possibilities.

Figure 15

Schematic Illustration of Three Possible Accounts Indicating How Biased Encoding and Attentional Refreshing may Explain the Value-Based Prioritization Effect


A
Possibility 1: Biased encoding only
Encoding approach:
(4)
(1)
(1)
(1)
Refreshing schedule:
Time
End of encoding:
Item strength
(4) (1) (1) (1)
End of maintenance:
Item strength
(4) (1) (1) (1)
B
Possibility 2: Biased attentional refreshing only
Encoding approach:
(4)
(1)
(1)
(1)
Refreshing schedule:
Time
OR:
Time
End of encoding:
Item strength
(4) (1) (1) (1)
End of maintenance:
Item strength
(4) (1) (1) (1)
C
Possibility 3: Biased encoding and refreshing
Encoding approach:
(4)
(1)
(1)
(1)
Refreshing schedule:
Time
OR:
Time
End of encoding:
Item strength
(4) (1) (1) (1)
End of maintenance:
Item strength
(4) (1) (1) (1)
Encoding approach interpretation:
Least time
Most time

Note: Panel A: A biased encoding only account. Participants prioritize the high value item during encoding, fixating on this item first and for longer periods of time relative to other items. Attentional refreshing is then applied in an equal way across items. Panel B: A biased attentional refreshing only account. Participants allocate attention equally across items during encoding, encoding each item with a similar strength. Participants then refresh the more valuable item more frequently or for longer periods of time. Panel C: A biased encoding and attentional refreshing account. Participants prioritize the high value item during encoding, fixating on this item first and for longer periods of time relative to other items. Participants then refresh the more valuable item more frequently or for longer periods of time. This account is best supported by current evidence (i.e. findings from the present study and Atkinson et al., 2022).

When considered alongside existing research, our findings also provide theoretical insights into differences (or similarities) between bottom-up, stimulus-driven attention and top-down, goal-directed attention in working memory. Previous research has found that bottom-up attentional manipulations (e.g. non-predictive cues) affect performance in change-detection tasks when information is presented simultaneously, but not during equivalent tasks that use sequential presentation (Ravizza et al., 2016). This led to the suggestion that the beneficial effects of bottom-up attention occur because participants encode the targeted item first, resulting in a primacy effect (Ravizza et al., 2016; Ravizza & Conn, 2022). Some experimental support for this notion has been provided (Ravizza et al., 2016). The current study builds on this, suggesting that the effects of value (which are considered to reflect a type of top-down attention; Allen et al., 2025) in working memory tasks are also partially explained by primacy. Taken together with Ravizza et al. (2016), this suggests that primacy at least partially explains both top-down and bottom-up attentional effects in working memory when information is presented simultaneously. Thus, overlapping mechanisms may drive these distinct forms of attentional direction in working memory tasks. It is, however, likely that a primacy-based account explains a greater proportion of bottom-up attentional effects, where individuals have little motivation to strategically prioritise the information.

More broadly, the present findings also have important implications regarding the mechanisms that determine which information is selected, maintained, and rendered consciously accessible in working memory. Recent iterations of high-level theoretical frameworks have adopted analogous approaches to outline how bottom-up capture and top-down control interact to influence selection of perceptual input in this way (Allen et al., 2026; Cowan et al., 2025). For example, the embedded processes perspective sets out how information maintained within the focus of attention serves as the basis for our ongoing conscious representation of internal thought processes and the world around us (Cowan et al., 2025). Within the multicomponent working memory model, the episodic buffer has been recharacterized as an awareness buffer, forming a modality-general and multi-source representational space in which information can be consciously maintained, manipulated, and prioritized (Allen et al., 2026). The present study demonstrates that specific patterns of selective attention, here indexed by eye fixation to items of varying value, help determine which representations are encoded and sustained within working memory, and thus which information enters a state of heightened accessibility and conscious awareness. In doing so, these findings provide support for the idea that moment-to-moment attentional allocation serves as a gatekeeper to the contents of working memory, influencing the prioritization of information (within the awareness buffer or focus of attention).

How robust are our findings? Would similar outcomes be observed if the encoding phase was shortened or lengthened? Our findings indicate that participants often fixate on high value item first, demonstrating that biases exist from the very start of the encoding phase. Thus, if the length of the encoding phase was further constrained, one would expect similar (or potentially even larger) biases towards high value items. Further, as our findings suggest that participants seek to preferentially encode the high value item, we would expect biases to persist even if the encoding phase was lengthened. As such, we believe participants would show a bias at encoding regardless of the length of the encoding phase. A related outstanding question of interest is whether the size of value-based prioritization effects would differ depending on the length of the retention interval. One possibility is that value-based prioritization effects would become larger as the length of the retention interval increases and participants have more opportunity to bias attentional refreshing towards particularly valuable information. This would result in a larger difference in recall error between high and lower value items with longer retention intervals. Alternatively, as the retention interval is lengthened and individuals have more time available for attentional refreshing, participants may more readily share the available time amongst items (though likely still preferentially refreshing the high value item to some extent). Based on such an explanation, value-based prioritization effects would still be expected to emerge with longer retention intervals, but may become smaller in size. It would be beneficial for further research to investigate these outstanding questions.

One boundary condition to the effects we describe here may be presentation method. Evidence that participants' pattern of fixations during the encoding phase differed as a function of value contrasts with a study by Ariel and Castel (2014), which examined value-directed remembering in a task that sequentially presented 12-item lists of words. They found no significant differences in fixation time for sequentially-presented low, medium, and high value words. These differences likely reflect the level of control that participants have during encoding. Although participants could have chosen not to fixate on particular items in Ariel and Castel (2014), they could not strategically choose which item to fixate on first or allocate a greater proportion of the encoding phase to high value items at the expense of low value items. As such, eye movements during item presentation may only drive value-based prioritization effects when information is presented simultaneously. Instead, individuals may engage in other strategies that seek to preferentially encode high value item information when sequential presentation is used. For example, individuals may refresh or spatially rehearse the higher value items during the interstimulus intervals between items. Alternatively, post-encoding processes such as attentional refreshing during maintenance may explain a greater proportion of the value effect when information is presented sequentially. We do not believe the establishment of such a boundary effect would significantly detract from the importance of our findings, as information is often encountered simultaneously in both research studies (e.g. Allen et al., 2012; Atkinson et al., 2022; Bays & Husain, 2008; Brockmole & Logie, 2013; Brown & Brockmole, 2010; Cowan et al., 2006; Forsberg et al., 2022; Hautekiet et al., 2025; Luck & Vogel, 1997; Oberauer et al., 2017; Park & Awh, 2026; Souza et al., 2014; Vogel et al., 2001; Zhang & Luck, 2008) and everyday tasks that rely on working memory. Moreover, the establishment of boundary conditions is important for both constraining and advancing working memory theory (Ueno & Allen, 2025). Evidence of distinct mechanisms underlying value effects for different presentation methods would also be in line with numerous studies highlighting key differences between simultaneous and sequential presentation (e.g. Allen et al., 2006; Atkinson et al., 2019; Bharti et al., 2020; Chen & Kuo, 2026; Magro et al., 2022).

To summarize, the current study provides novel evidence that participants vary their eye movements during working memory tasks as a function of the notional value of items. This enables high value information to be encoded first (thus benefiting from primacy), as well as more strongly relative to other items. Engagement in such strategies partially explains value-based prioritization effects in working memory. These findings provide strong evidence that maintenance mechanisms are not the only factor that contributes to value-based prioritization effects, establishing an important role of encoding processes in determining which information is more likely to remain at least briefly consciously accessible in the service of task performance. Further research should investigate how distinct the operations of these processes are, their relative importance across different task contexts, and whether strategies at other stages (e.g. retrieval) also partially drive value-based prioritization effects.

References

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, *135*(2), 298-313. 10.1037/0096-3445.135.2.298

Allen, R. J., Hitch, G. J., Mate, J., & Baddeley, A. D. (2012). Feature binding and attention in working memory: A resolution of previous contradictory findings. *Quarterly Journal of Experimental Psychology*, *65*(12), 2369-2383. 10.1080/17470218.2012.687384

Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention, Perception, & Psychophysics*, *80*, 1731-1743. 10.3758/s13414-018-1543-6.

Allen, R. J., Atkinson, A. L., & Nicholls, L. A. B. (2021). Strategic prioritisation enhances young and older adults' visual feature binding in working memory. *Quarterly Journal of Experimental Psychology*, *74*(2), 363-376. 10.1177/1747021820960712.

Allen, R. J., Atkinson, A., & Hitch, G. J. (2025). Getting value out of working memory through strategic prioritisation; implications for storage and control. *Quarterly Journal of Experimental Psychology, 7*8(2):405-424. 10.1177/17470218241258102.

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2017). Executive and perceptual distraction in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(9), 1677-1693. 10.1037/xhp0000413.

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2026). Awareness as the heart of working memory. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2026.04.001

Allen, R., & Atkinson, A. (2021). Retrospective and prospective prioritization in visual working memory. Preprint. Available at: https://osf.io/4x8zu_v1

Ariel, R., & Castel, A. D. (2014). Eyes wide open: Enhanced pupil dilation when selectively studying important information. *Experimental Brain Research*, *232*(1), 337-344. 10.1007/s00221-013-3744-5.

Atkinson, A. L., Allen, R. J., Baddeley, A. D., Hitch, G. J., & Waterman, A. H. (2021). Can valuable information be prioritized in verbal working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(5), 747-764. 10.1037/xlm0000979

Atkinson, A. L., Berry, E. D., Waterman, A. H., Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2018). Are there multiple ways to direct attention in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 115-126. 10.1111/nyas.13634

Atkinson, A. L., Oberauer, K., Allen, R. J., & Souza, A. S. (2022). Why does the probe value effect emerge in working memory? Examining the biased attentional refreshing account. *Psychonomic Bulletin & Review*, *29*(3), 891-900. 10.3758/s13423-022-02056-6

Atkinson, A. L., Pinheiro Sanchez, B., Warburton, M., Allmark, H., & Allen, R. J. (2025). The Ability to Direct Attention in Working Memory Is Not Impaired in Adults With Symptoms of ADHD. *Journal of Attention Disorders*, 29(9), 684-705. 10.1177/10870547251330039

Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2019). Can children prioritize more valuable information in working memory? An exploration into the effects of motivation and memory load. *Developmental Psychology*, *55*(5), 967-980. 10.1037/dev0000692

Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2024). Does value-based prioritization at working memory enhance long-term memory? *Memory & Cognition*, 1-16. 10.3758/s13421-024-01532-9

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119-126. 10.1016/s1364-6613(00)01593-x

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2021). A multicomponent model of working memory. In R. H. Logie, V. Camos & N. Cowan (Eds.), Working memory: State of the science (pp. 10–43). Oxford University Press.

Barrouillet, P., & Camos, V. (2021). The time-based resource-sharing model of working memory. *Working memory: State of the science*, 85-115.

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 7. 1–11. 10.1167/9.10.7

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*(5890), 851-854. 10.1126/science.1158023

Bharti, A. K., Yadav, S. K., & Jaswal, S. (2020). Feature binding of sequentially presented stimuli in visual working memory. *Frontiers in Psychology*, *11*, 33. 10.3389/fpsyg.2020.00033

Beesley, T & Ivory, M. (2025). Eyetools. Tools for eye data analysis. https://tombeesley.github.io/eyetools/

Bhanap, R., Oberauer, K., & Rosner, A. (2025). Investigating retrieval strategies in an associative recognition test in working memory: Evidence from eye movements. *Cognition*, *263*, 106199.

Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55,753 participants aged 8–75. *Frontiers in Psychology*, *4*, 12. 10.3389/fpsyg.2013.00012

Brown, L. A., & Brockmole, J. R. (2010). The role of attention in binding visual features in working memory: Evidence from cognitive ageing. *Quarterly Journal of Experimental Psychology*, *63*(10), 2067-2079. 10.1080/17470211003721675

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: the executive and phonological loops. *Frontiers in Human Neuroscience*, *8*, 900. 10.3389/fnhum.2014.00900

Camos, V., Johnson, M., Loaiza, V., Portrat, S., Souza, A., & Vergauwe, E. (2018). What is attentional refreshing in working memory? *Annals of the New York Academy of Sciences*, *1424*(1), 19-32. 10.1111/nyas.13616

Camos, V., Lagner, P., & Barrouillet, P. (2009). Two maintenance mechanisms of verbal information in working memory. *Journal of Memory and Language*, *61*(3), 457-469. 10.1016/j.jml.2009.06.002

Chen, Y. T., & Kuo, B. C. (2026). Simultaneous and Sequential Presentations Differentially Modulate the Temporal Dynamics of Working Memory Processes. *Journal of Cognitive Neuroscience*, *38*(4), 680-694. 10.1162/JOCN.a.2399.

Cinar, H., Atkinson, A. L., Waterman, A. H., & Allen, R. J. (2025). Prioritizing feature bindings across space and modality in working memory. *Memory & Cognition*, 1-17.

Cowan, N (1999). An Embedded-Processes Model of working memory. In: A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). New York, US: Cambridge University Press.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*, 197-223. 10.1007/s10648-013-9246-y

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*, 1158-1170. 10.3758/s13423-016-1191-6

Cowan, N., Ahmed, N. I., Bao, C., Cissne, M. N., Yee-Flores, R. D., Gutierrez, R. M., ... & Superbia-Guimarães, L. (2025). Theories of consciousness from the perspective of an embedded processes view. *Psychological Review*, *132*(1), 76-106.

Cowan, N., Morey, C. C., & Naveh-Benjamin, M. (2021). An embedded-processes approach to working memory. *Working memory: State of the science*, 44-84.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult?. *Developmental Psychology*, *42*(6), 1089-102. 10.1037/0012-1649.42.6.1089.

Cui, J. F., Wang, Y., Shi, H. S., Liu, L. L., Chen, X. J., & Chen, Y. H. (2015). Effects of working memory load on uncertain decision-making: Evidence from the Iowa Gambling Task. *Frontiers in Psychology*, *6*, 162. 10.3389/fpsyg.2015.00162

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19-23. 10.1111/1467-8721.00160

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. 10.3758/BF03193146

Forsberg, A., Guitard, D., Adams, E. J., Pattanakul, D., & Cowan, N. (2022). Children's long-term retention is directly constrained by their working memory capacity limitations. *Developmental Science*, *25*(2), e13164. 10.1111/desc.13164

Frischkorn, G. T., & Popov, V. (2025). A tutorial for estimating Bayesian hierarchical mixture models for visual working memory tasks: Introducing the Bayesian Measurement Modeling (bmm) package for R. *Behavior Research Methods*, *57*(5), 144. 10.3758/s13428-025-02643-0

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. 10.1214/ss/1177011136

Hautekiet, C., Langerock, N., & Vergauwe, E. (2024). Prioritization in visual working memory: An investigation of distractor susceptibility and different prioritization modes. Preprint. Available at: https://osf.io/ae9y4

Hautekiet, C., Langerock, N., & Vergauwe, E. (2025). Prioritization in visual working memory: An investigation of distractor susceptibility and different prioritization modes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 10.1037/xlm0001534

Hitch, G. J., Allen, R. J., & Baddeley, A. D. (2020). Attention and binding in visual working memory: Two forms of attention and two kinds of buffer storage. *Attention, Perception, & Psychophysics*, *82*, 280-293. 10.3758/s13414-019-01837-x

Hitch, G. J., Allen, R. J., & Baddeley, A. D. (2025). The multicomponent model of working memory fifty years on. *Quarterly Journal of Experimental Psychology*, *78*(2), 222-239. 10.1177/17470218241290909

Hitch, G. J., Hu, Y., Allen, R. J., & Baddeley, A. D. (2018). Competition for the focus of attention in visual working memory: perceptual recency versus executive control. *Annals of the New York Academy of Sciences*, *1424*(1), 64-75. 10.1111/nyas.13631

Hu, Y., Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2016). Executive control of stimulus-driven and goal-directed attention in visual working memory. *Attention, Perception, & Psychophysics*, *78*, 2164-2175. 10.3758/s13414-016-1106-7

Hu, Y., Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2023). Visual working memory phenomena based on categorical tasks replicate using a continuous measure: A simple interpretation and some methodological considerations. *Attention, Perception, & Psychophysics*, *85*(5), 1733-1745. 10.3758/s13414-023-02656-x

Hu, Y., Hitch, G. J., Baddeley, A. D., Zhang, M., & Allen, R. J. (2014). Executive and perceptual attention play different roles in visual working memory: evidence from suffix and strategy effects. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1665-1678. 10.1037/a0037163

Jeanneret, S., Bartsch, L. M., & Vergauwe, E. (2023). To be or not to be relevant: Comparing short-and long-term consequences across working memory prioritization procedures. *Attention, Perception, & Psychophysics*, *85*(5), 1486-1498. 10.3758/s13414-023-02706-4

Kelter, R. (2020). Analysis of Bayesian posterior significance and effect size indices for the two-sample t-test to support reproducible medical research. *BMC Medical Research Methodology*, *20*, 1-18.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389-433. 10.1016/S0160-2896(05)80012-1

Lemaire, B., Pageot, A., Plancher, G., & Portrat, S. (2018). What is the time course of working memory attentional refreshing?. *Psychonomic Bulletin & Review*, *25*(1), 370-385. 10.3758/s13423-017-1282-z

Lenth R (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package. https://CRAN.R-project.org/package=emmeans

Loaiza, V. M., & Souza, A. S. (2022). The eyes don't have it: Eye movements are unlikely to reflect refreshing in working memory. *PloS one*, *17*(7), e0271116. 10.1371/journal.pone.0271116

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279-281. 10.1038/36846

Magro, L., Mirault, J., Grainger, J., & Majerus, S. (2022). Sequential versus simultaneous presentation of memoranda in verbal working memory:(How) does it matter? *Memory & Cognition*, *50*(8), 1756-1771. 10.3758/s13421-022-01284-4

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software, 4*(40), 1541. doi:10.21105/joss.01541

Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: protection or prioritization? *Perception & Psychophysics*, *69*(8), 1422-1434.

Mora, G., & Camos, V. (2015). Dissociating rehearsal and refreshing in the maintenance of verbal information in 8-year-old children. *Frontiers in Psychology*, *6*, 11. 10.3389/fpsyg.2015.00011

Morey, C. C., Mareva, S., Lelonkiewicz, J. R., & Chevalier, N. (2018). Gaze-based rehearsal in children under 7: A developmental investigation of eye movements during a serial spatial memory task. *Developmental Science*, *21*(3), e12559.

Morey. R., & Rouder, J. (2022). BayesFactor: Computation of Bayes Factors for Common Designs. https://CRAN.R-project.org/package=BayesFactor

Nikolov, T. Y., Allen, R. J., Havelka, J., Darling, S., van de Vegte, B., & Morey, C. C. (2025). Navigating the mind's eye: Understanding gaze shifts in visuospatial bootstrapping. *Quarterly Journal of Experimental Psychology*, *78*(2), 391-404. 10.1177/17470218241282426

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(3),* 411-421. 10.1037/0278-7393.28.3.411

Oberauer, K. (2013). The focus of attention in working memory—from metaphors to mechanisms. *Frontiers in Human Neuroscience, 7,* 673. 10.3389/fnhum.2013.00673

Oberauer, K. (2019). Working memory and attention–A conceptual analysis and review. *Journal of Cognition*, *2*(1), 36. 10.5334/joc.58

Oberauer, K., Awh, E., & Sutterer, D. W. (2017). The role of long-term memory in a test of visual working memory: Proactive facilitation but no proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(1), 1-22. 10.1037/xlm0000302

Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, *21*(3), 164-169. 10.1177/0963721412444727

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*(5), 779-819. 10.3758/s13423-012-0272-4

Park, H. B., & Awh, E. (2026). Contributions from long-term memory explain superior visual working memory performance with meaningful objects. *Journal of Neuroscience*, *46*(5). 10.1523/JNEUROSCI.1660-25.2025

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, *51*, 195-203.

Ravizza, S. M., & Conn, K. M. (2022). Gotcha: Working memory prioritization from automatic attentional biases. *Psychonomic Bulletin & Review*, *29*(2), 415-429. 10.3758/s13423-021-01958-1

Ravizza, S. M., Uitvlugt, M. G., & Hazeltine, E. (2016). Where to start? Bottom-up attention improves working memory by determining encoding order. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(12), 1959. 10.1037/xhp0000275

Roe, D., Allen, R. J., Elsley, J., Miles, C., & Johnson, A. J. (2024). Working memory prioritisation effects in tactile immediate serial recall. *Quarterly Journal of Experimental Psychology*, 17470218241231283. 10.1177/17470218241231283

Sandry, J., & Ricker, T. J. (2020). Prioritization within visual working memory reflects a flexible focus of attention. *Attention, Perception, & Psychophysics*, *82*, 2985-3004. 10.3758/s13414-020-02049-4

Sandry, J., Schwark, J. D., & MacDonald, J. (2014). Flexibility within working memory and the focus of attention for sequential verbal information does not depend on active maintenance. *Memory & Cognition*, *42*, 1130-1142. 10.3758/s13421-014-0422-1

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. (2023). afex: Analysis of Factorial Experiments. https://CRAN.R-project.org/package=afex.

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics*, *78*, 1839-1860. 10.3758/s13414-016-1108-5

Souza, A. S., Czoschke, S., & Lange, E. B. (2020). Gaze-based and attention-based rehearsal in spatial working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(5), 980-1003. 10.1037/xlm0000771

Souza, A. S., Rerko, L., Lin, H. Y., & Oberauer, K. (2014). Focused attention improves working memory: implications for flexible-resource and discrete-capacity models. *Attention, Perception, & Psychophysics*, *76*(7), 2080-2102. 10.3758/s13414-014-0687-2

Souza, A. S., Rerko, L., & Oberauer, K. (2015). Refreshing memory traces: Thinking of an item improves retrieval from visual working memory. *Annals of the New York Academy of Sciences*, *1339*(1), 20-31. 10.1111/nyas.12603

Souza, A. S., Vergauwe, E., & Oberauer, K. (2018). Where to attend next: guiding refreshing of visual, spatial, and verbal representations in working memory. *Annals of the New York Academy of Sciences*, *1424*(1), 76-90. 10.1111/nyas.13621

Süß, H. M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence*, *30*(3), 261-288. 10.1016/S0160-2896(01)00100-3

Tingley, D., Yamamoto T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software, 59*(5), 1-38. http://www.jstatsoft.org/v59/i05/.

Ueno, T., & Allen, R. J. (2025). Running after two hares in visual working memory: Exploring retrospective attention to multiple items using simulation, behavioral outcomes, and eye tracking. *Journal of Experimental Psychology: Human Perception and Performance, 51*(3), 405–420. 10.1037/xhp0001270

Vergauwe, E., Hautekiet, C., Zhang, Z., & Lewis-Peacock, J. (2025). The consequences of strategic prioritization in working memory. Preprint. Available at: https://osf.io/4a2bu.

Vergauwe, E., Souza, A. S., Langerock, N., & Oberauer, K. (2024). The effect of instructed refreshing on working memory: Is the memory boost a function of refreshing frequency or refreshing duration? *Memory & Cognition*, 1-13. 10.3758/s13421-024-01666-w

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92-114. 10.1037//0096-1523.27.1.92

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2005). Pushing around the locus of selection: Evidence for the flexible-selection hypothesis. *Journal of Cognitive Neuroscience*, *17*(12), 1907-1922. 10.1162/089892905775008599

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235. 10.1038/nature06860

Supplementary materials

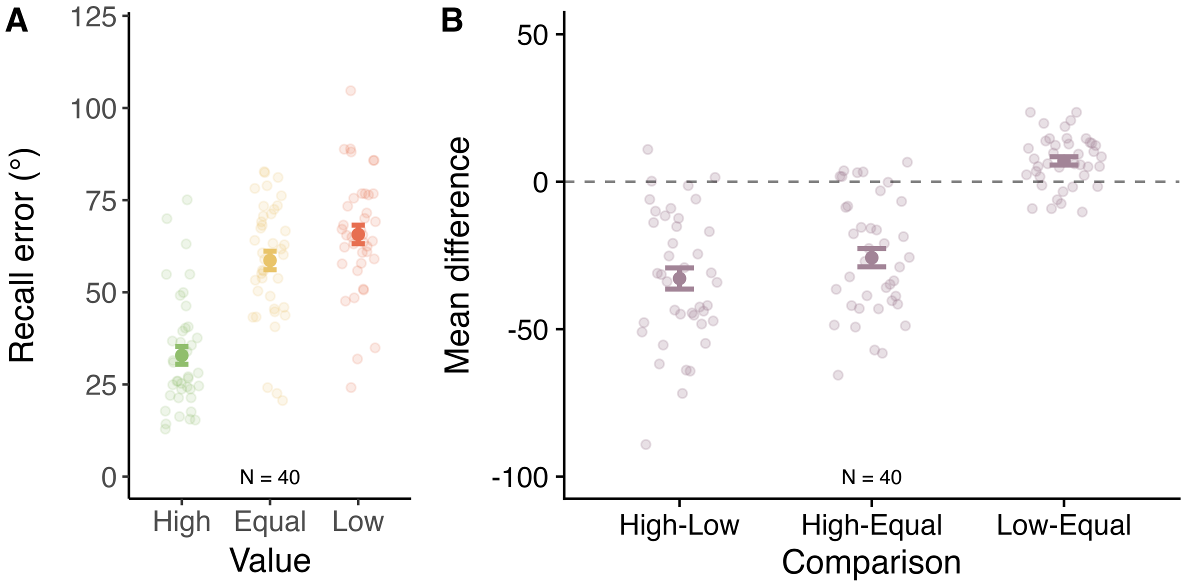
1. Experiment 1: Behavioural outcomes for all participants

This analysis includes all participants with usable behavioural data ($N = 40$). This excludes one participant due to a technical error, but includes the two participants excluded from the analysis in the main body due to poor eye-tracking. Mean recall error is displayed in Figure S.1.1.A, as a function of value. Figure S.1.1.B displays the difference in recall error between the value conditions. A one-way repeated measures ANOVA was conducted to investigate the effect of value on recall error. There was a significant effect of value (*Greenhouse-Geisser (GG) corrected* $F(1.23, 47.96) = 73.12$, $MSE = 265.48$, $p < .001$, $\eta_p^2 = .65$; $BF_{10} >$ 10,000). Post-hoc tests revealed that recall error was significantly lower for high value items ($M = 32.9$, $SE = 2.41$) relative to equal ($M = 58.6$, $SE = 2.51$; $p < .001$, $d = -1.31$; $BF_{10} >$ 10,000) and low value ($M = 65.7$, $SE = 2.51$; $p < .001$, $d = -1.45$; $BF_{10} >$ 10,000) items. Recall error was also significantly lower for equal value items relative to low value items ($p < .001$, $d = -0.79$; $BF_{10} = 1682.78$).

Therefore, the conclusions from this analysis did not differ from those presented in the main body.

Figure S.1.1

Task Performance in Experiment 1 for all Participants with Usable Task Data (N = 40)



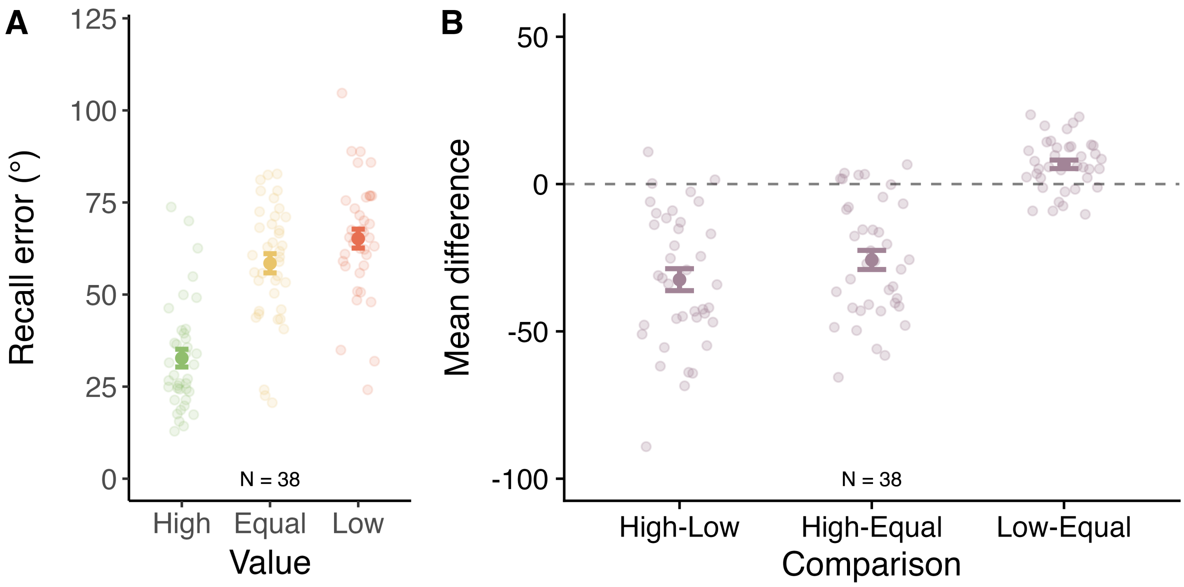
Note. Panel A displays mean recall error, as a function of value (higher recall error reflects worse performance). Panel B displays the mean difference between the value conditions, calculated as the first condition listed minus the second (e.g. the figure for High vs Low for a given participant was calculated by subtracting their mean recall error in the “Low” value condition from their mean recall error in the “High” condition). Across panels, bold points reflect mean across participants, whilst error bars denote standard error. Lighter circles in the background indicate individual participants’ mean.

2. **Experiment 1: Behavioural outcomes for only trials with fixations recorded for participants analysed in main body (*N* = 38).**

This analysis reports behavioural outcomes for participants reported in the main body, but focusing only on trials with usable eye-tracking data (N = 38). Mean recall error as a function of value is displayed in Figure S.2.1.A, whilst the mean difference in recall error between conditions is displayed in Figure S.2.1.B. A one-way repeated measures ANOVA revealed a significant effect of value (*Greenhouse-Geisser (GG) corrected* $F(1.23, 45.36) = 66.39$, $MSE = 274.06$, $p < .001$, $\eta_p^2 = .64$; $BF_{10} > 10{,}000$). Post-hoc tests revealed lower recall error in the high value condition ($M = 32.7$, $SE = 2.41$) relative to equal ($M = 58.5$, $SE = 2.60$; $p < .001$, $d = -1.29$; $BF_{10} > 10{,}000$) and low value conditions ($M = 65.2$, $SE = 2.58$; $p < .001$, $d = -1.41$; $BF_{10} > 10{,}000$). Moreover, recall error was lower for equal value items relative to low value items ($p < .001$, $d = -0.75$; $BF_{10} = 522.80$). Thus, the conclusions from this analysis did not differ from those presented in the main body.

Figure S.2.1

Task Performance in Experiment 1 for Only Trials With Eye-Tracking Fixations Recorded During the Encoding Phase (N = 38)



Note. Panel A displays mean recall error, as a function of value (higher recall error reflects worse performance). Panel B displays the mean difference between the value conditions, calculated as the first condition listed minus the second (e.g. the figure for High vs Low for a given participant was calculated by subtracting their mean recall error in the “Low” value condition from their mean recall error in the "High" condition). Across panels, bold points reflect mean across participants, whilst error bars denote standard error. Lighter circles in the background indicate individual participants' mean.

3. Experiment 1 response time analyses (*N* = 38)

We conducted analyses of the response times to ensure that the effects of value on recall error were not due to a speed-accuracy trade off. We conducted this on data for all trials for participants included in the final sample.

Pre-processing

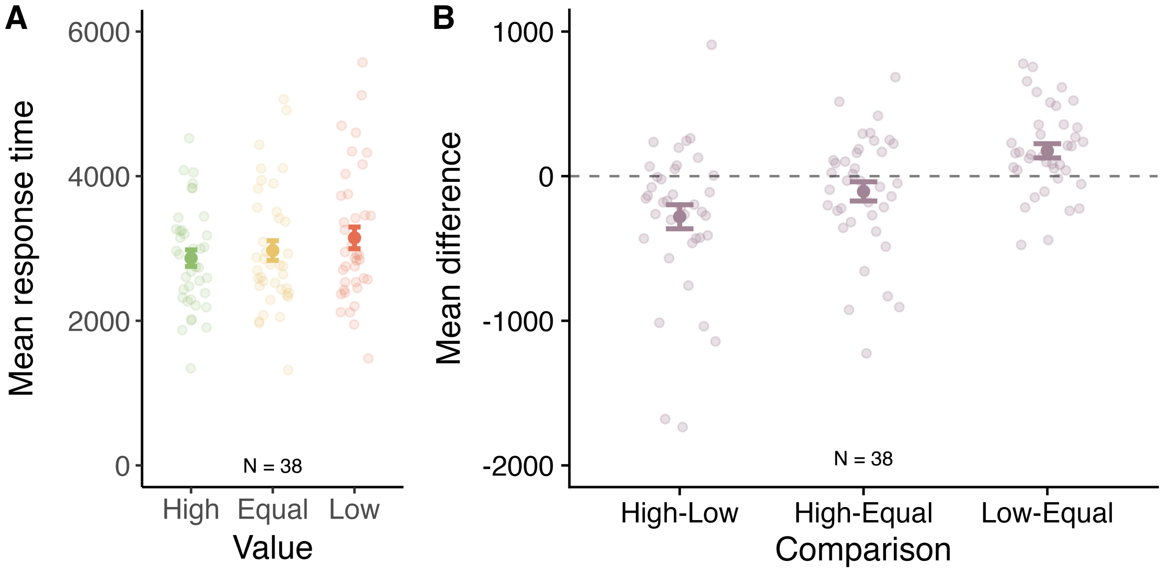
We first excluded trials where the response time was over 20,000ms (20 seconds; one trial excluded). We then excluded trials that fell 2.5 SDs above or below the mean for each value condition for each participant.

Results

Mean response times as a function of value are displayed in Figure S.3.1.A, whilst Figure S.3.1.B displays the difference in response time between the three value conditions. A one-way repeated measures ANOVA was conducted to investigate the effect of value (three levels: High, Equal, Low) on response times. There was a significant effect of value (GG-corrected, $F(1.51, 55.86) = 8.78$, $MSE = 115384.18$, $p = .001$, $\eta_p^2 = .19$; $BF_{10} = 68.80$). Post-hoc tests revealed significantly slower response times for low value items ($M = 3147$, $SE = 150$) relative to high value ($M = 2866$, $SE = 115$; $p = .004$, $d = -0.55$; $BF_{10} = 18.89$) and equal value ($M = 2971$, $SE = 136$; $p = .003$, $d = -0.58$; $BF_{10} = 30.77$, $BF_{01} = 1.82$) items. There was no significant difference between high and equal value items ($p = .121$, $d = -0.26$, $BF_{10} = 0.55$).

Figure S.3.1

Response Times in Experiment 1



Note. Panel A displays mean response times, as a function of value. Panel B displays the mean difference between the three value conditions. The dotted line at 0 reflects no difference. Values above 0 indicate slower response times in the first condition listed, whilst values below 0 indicate faster response times in the first condition listed. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error (SE) of the mean.

Therefore, although there was a significant effect of value on response times, this did not reflect slower response times for high value items. In fact, high value items were responded to significantly faster than low value items, and numerically faster than equal value items.

4. Experiment 1: Analysis of the proportion of fixations and length of individual fixations

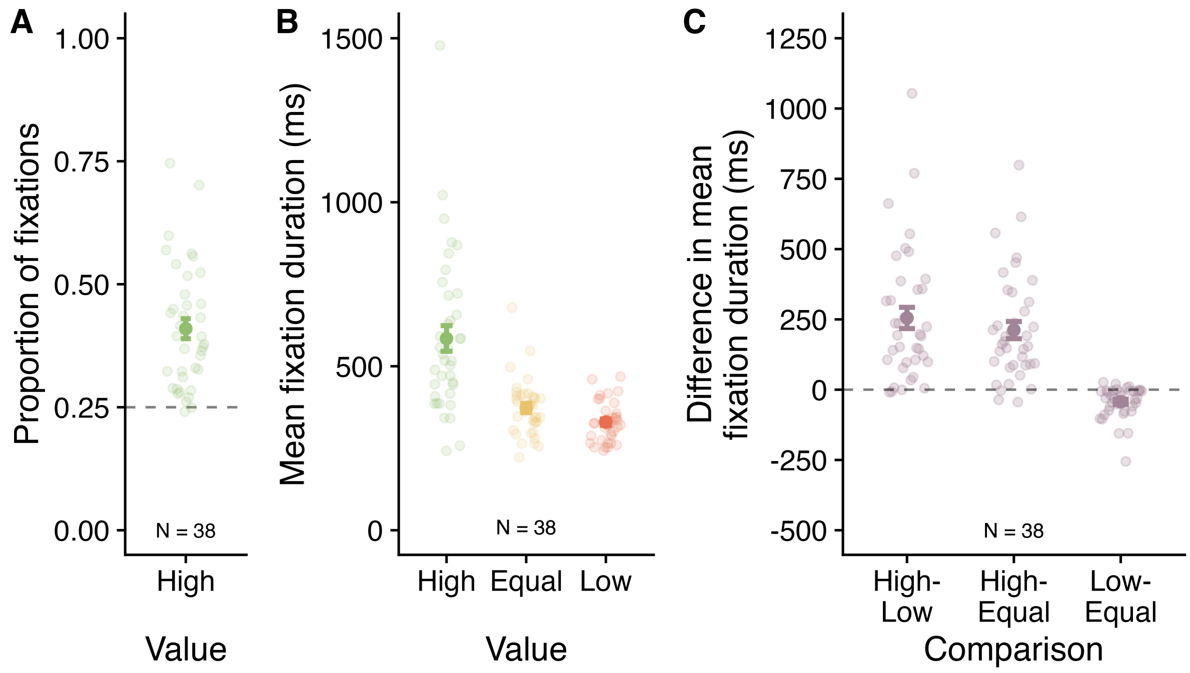
As shown in the main analyses, total time spent fixating on High value items during the encoding phase is greater than that for Equal and Low value items. This effect could be due to a greater number of fixations towards high value items, longer fixations on high value items, or a combination of both. This was examined in subsidiary analyses.

Proportion of fixations

We first assessed the proportion of fixations (a measure that ignores the length of those fixations) that participants made towards high value items, and whether this was greater than what would be expected based on an equal distribution of fixations across the four items in the array (0.25). In this analysis, the proportion was calculated by dividing the number of fixations participants made to high value items by the total number of fixations participants made to all items during the encoding phase of a given trial (excluding fixations that did not fall on an item). Trials were excluded where participants did not fixate on any of the items during the encoding phase. The mean proportion of fixations towards high value items during the encoding phase is displayed in Figure S.4.1A. A one sample t-test indicated that the mean proportion of fixations towards high value items (M = 0.41, SE = .02) was significantly different from that expected based on an equal distribution of fixations across the four items (0.25), $t(37) = 7.84$, $p < .001$, $d = 1.27$).

Figure S.4.1

Other Eye-Tracking Outcomes in Experiment 1



Note. Panel A: Mean proportions of fixations towards the high value item in the encoding phase of the differential value trials. The dotted line at 0.25 reflects an equal distribution of fixations across the four items. Panel B shows the mean fixation duration (in ms), as a function of item value. Panel C displays the difference in mean duration between the value conditions. The dotted line at 0 reflects no difference. Values above 0 indicate that participants' individual fixations were longer for the first condition listed relative to the second. Meanwhile, values below 0 indicate that participants' individual fixations were shorter for the first condition listed relative to the second. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error of the mean.

Length of individual fixations

We next assessed the mean length of individual fixations as a function of item value. For each trial, we determined the mean fixation duration for each value condition. We then averaged across trials for each participant. These data reflected valid fixations on items only, and hence data was not scored as 0 if the participant failed to fixate an item. This approach (as opposed to scoring a 0 where participants did not fixate on an item at all) ensures this analysis is not confounded by the difference in the total number or proportion of fixations (as detailed above).

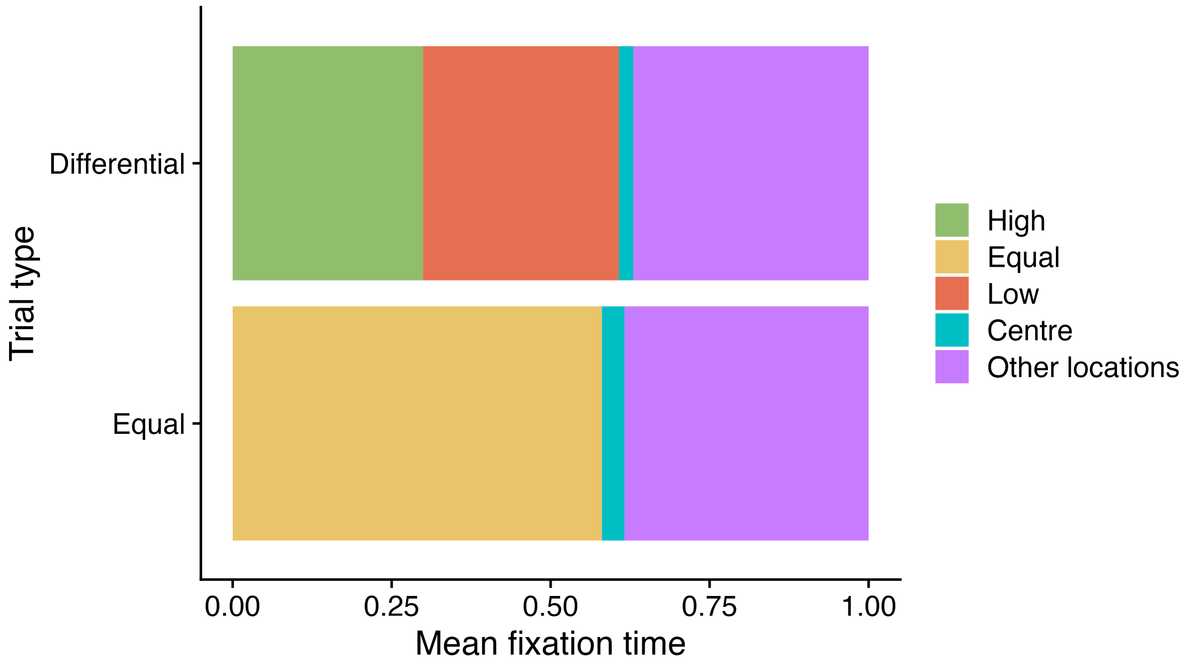
Mean duration of each fixation is displayed in Figure S.4.1.B, as a function of value. A one-way repeated measures ANOVA was conducted with value (three levels: High, Equal, Low) as the independent variable and fixation length as the dependent variable. This revealed a significant effect of value (*GG-corrected F*(1.04, 38.42) = 45.58, *MSE* = 29927.90, $p < .001$, $\eta_p^2 = .55$; $BF_{10} > 10{,}000$). The length of fixations was significantly longer for high value items ($M = 585$, $SE = 39.09$) relative to equal ($M = 373$, $SE = 13.78$; $p < .001$; $d = 1.12$; $BF_{10} > 10{,}000$) and low value items ($M = 330$, $SE = 9.54$; $p < .001$, $d = 1.09$; $BF_{10} > 10{,}000$). Fixations towards equal value items were also longer than fixations towards low value items ($p < .001$; $d = 0.76$; $BF_{10} = 615.80$).

5. Fixations including screen centre as an area of interest (AOI; Experiment 1)

Figure S.5.1 shows the mean proportion of the encoding phase participants spent fixating on the spatial location of high value items, low value items, equal value items, the screen centre[9], and other areas of the screen.

Figure S.5.1

Mean Proportion of the Encoding Phase Participants Spent Fixating on High Value Items, Low Value Items, Equal Value Items, the Screen Centre, and Other Areas of the Screen



Note. Data is displayed separately for differential value trials (where one item was high value and the rest low value) and equal value trials (where all items were of equal value). Bars show the sum across that item value. Thus, green bars reflect fixations to one high value item per trial, red bars reflect fixations towards three low value items per trial, and yellow bars depict fixations to four equal value items per trial.

[9] The area of interest (AOI) was 350x350 pixels around the screen centre.

6. Encoding fixation duration as a mediator (further analysis; Experiment 1)

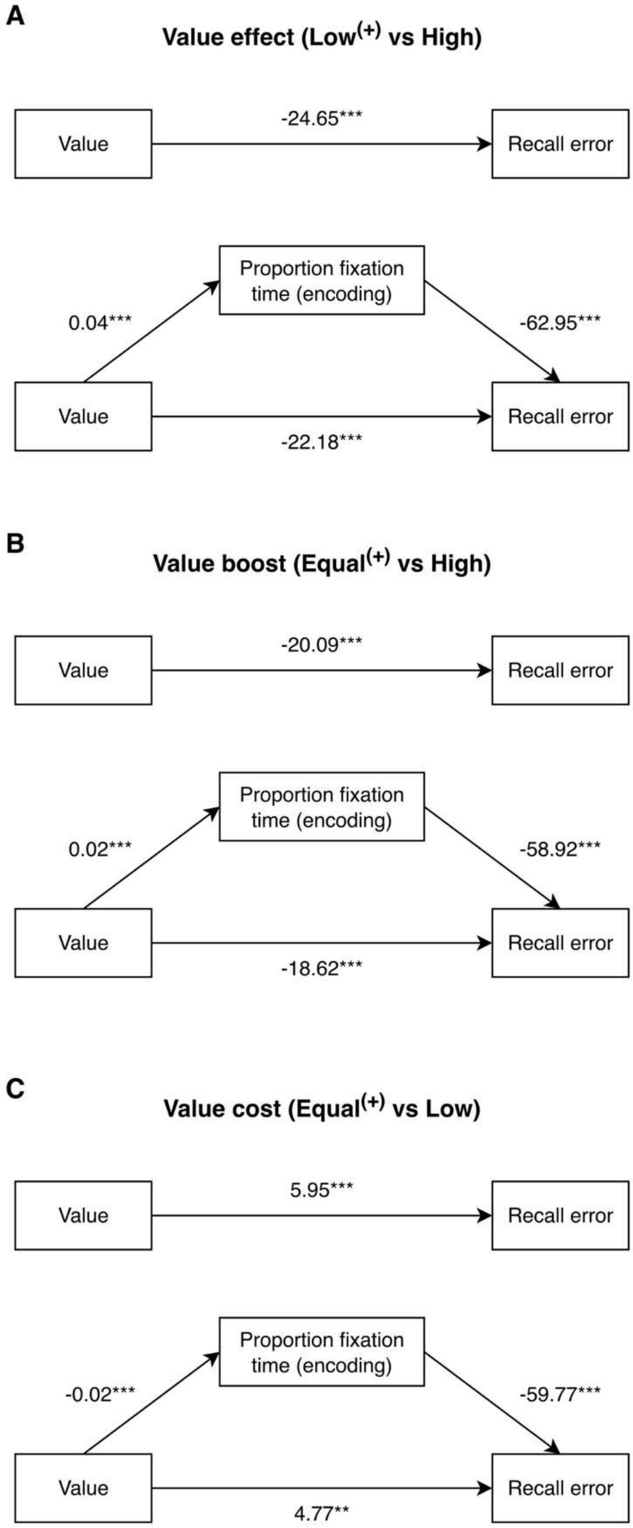
A series of mediation models were conducted to investigate whether encoding fixation duration still partially mediated the effect of value on recall error after trials were excluded whether the first fixation was on the tested item. This examines whether encoding fixation duration still mediates the relationship after accounting for a primacy-based effect.

As in the main analysis, separate mediation models were conducted for the overall value effect (low vs high value), the value boost (equal vs high value), and the value cost (equal vs low value). Outcomes are displayed in Figure S.6.1. In all models, the effect of value on recall error was partially mediated by the proportion of time participants fixated on the tested item during encoding. The indirect effects were -2.48 for the overall value effect (low vs high; 95% CI: -3.41--1.69; $p < .001$), -1.47 for the value boost (equal vs high; 95% CI: -2.54--0.63; $p < .001$), and 1.19 for the value cost (equal vs low; 95% CI: 0.76-1.68; $p < .001$). The proportion of the effect mediated was 0.10 (95% CI: 0.07-0.15; $p < .001$) for the overall value effect (low vs high), 0.07 (95% CI: 0.03-0.13; $p < .001$) for the value boost (equal vs high), and 0.20 (95% CI: 0.11-0.45; $p < .001$) for the value cost (equal vs low).

These outcomes indicate the effect of value on recall error is partially mediated by the proportion of time participants spend fixating on an item during encoding, even after when removing trials where the first item fixated on was the tested item. This indicates the mediation cannot be fully attributed to a primacy-like account.

Figure S.6.1

Outcomes from the Mediation Analysis in Experiment 1 Examining Encoding Fixation Duration as a Mediator, after Excluding Trials where the First Fixation was on the Tested Item



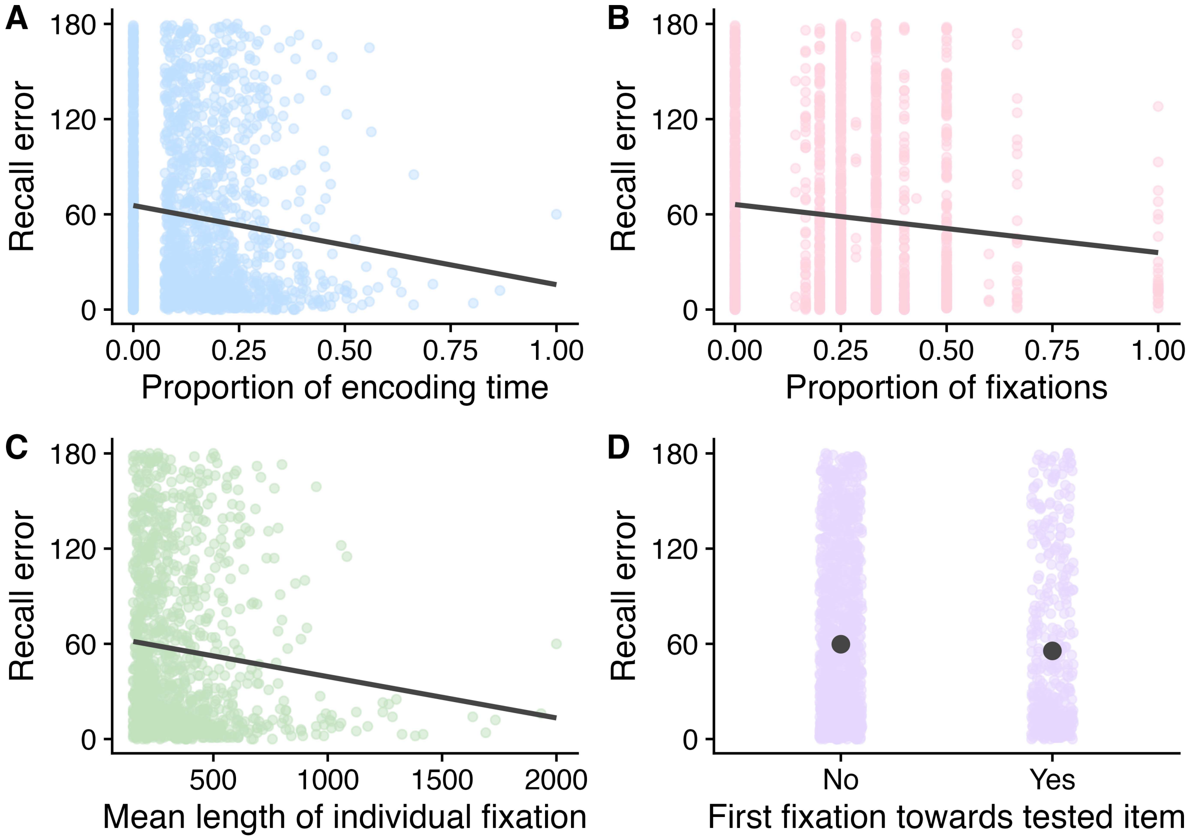
Note. Panel A reflects the overall value effect (low vs high value), Panel B reflects the value boost (equal vs high value), and Panel C reflects the value cost (equal vs low value). In all panels, the upper diagram reflects the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. $^{+}$ denotes reference category. *** $p < .001$, ** $p < .01$, * $p < .05$.

7. Outcomes for analyses examining whether eye movement measures predict recall error in equal value trials (Experiment 1)

The association between the eye tracking parameters during encoding and recall error in the equal value trials is presented in Figure S.7.1. Analysis was conducted to investigate whether the main eye tracking outcomes analysed across the experiments predicted recall error. A separate mixed-effect model was fit for each eye-tracking measure, with the eye-tracking measure as the predictor, recall error as the outcome, and participant ID as a random intercept.

Figure S.7.1

Relationship Between the Eye-Tracking Parameters during Encoding and Recall Error in the Equal Value Trials of Experiment 1



Note. Panel A shows the proportion of encoding time. Panel B displays the proportion of fixations towards the tested item, Panel C shows mean length of individual fixations, and Panel D displays the first fixation towards tested item (no/yes). In panel D, the black circles shows the mean recall error.

Encoding fixation duration: Figure S.7.1.A shows the association between the proportion of encoding time participants spent fixating on the tested item and recall error. The proportion of time participants spent fixating on the tested item during the encoding phase significantly predicted recall error (β = -54.95, SE = 10.19, t = -5.39, p < .001), with a higher proportion associated with lower recall error.

Proportion of fixations towards the tested item: Figure S.7.1.B shows the relation between the proportion of fixations towards the tested item and recall error. The proportion of fixations significantly predicted recall error (β = -32.51, SE = 6.61, t = -4.92, p < .001), with a higher value associated with lower recall error.

Length of individual fixations: This was calculated as in the main analyses. Figure S.7.1.C displays the association between the length of individual fixations towards the tested item and recall error. The length of individual fixations towards the tested item significantly predicted recall error (β = -0.03, SE = 0.01, t = -3.76, p < .001), with longer fixations associated with lower recall error.

First fixation: Figure S.7.1.D displays recall error as a function of whether the tested item was fixated on first or not. Whether the first fixation was towards the tested item did not significantly predict recall error, though it did approach significance (β = -5.32, SE = 2.96, t = -1.80, p = .073).

8. Experiment 1: Maintenance fixation duration

Analysis was performed to investigate whether the proportion of maintenance time participants spent fixating on locations in which items had appeared differs as a function of value; and whether this factor mediates the relationship between value and recall error.

Data processing

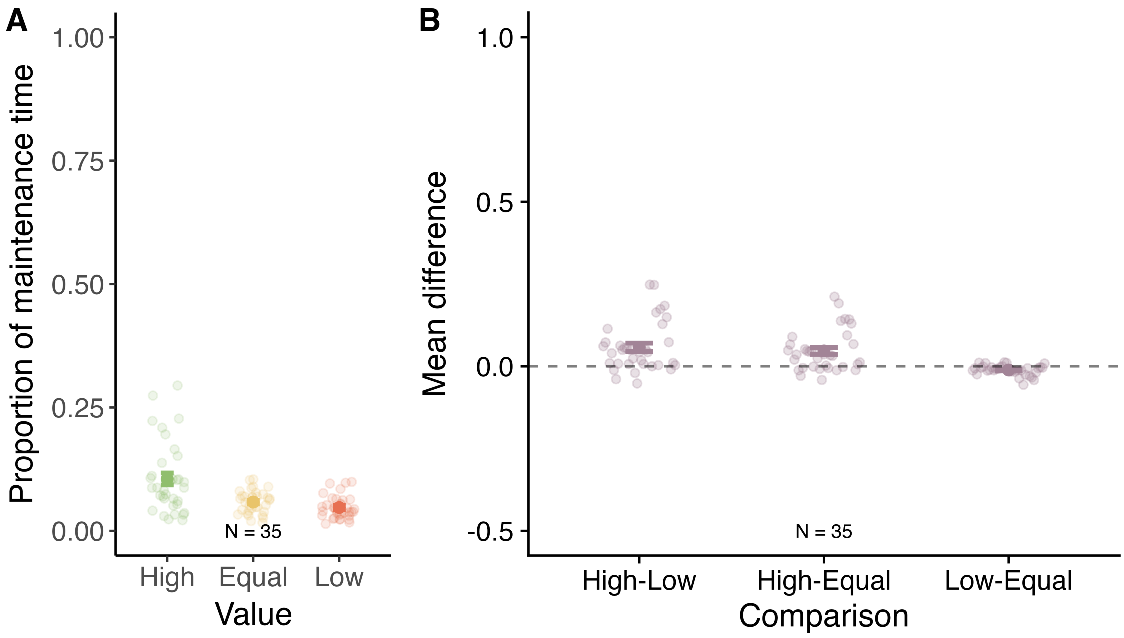
In addition to the participants excluded in the main analysis (see *Participant* section of Experiment 1), a further three participants were excluded as the number of trials with no fixations during maintenance was over 2.5 standard deviations above the mean. The analysis was therefore performed on data for 35 participants. In the analysed sample, the percentage of trials with no recorded fixations was low (M = 0.21, SD = 0.70). Trials with no recorded fixations were excluded. Following this, the mean number of fixations across participants in the maintenance phase was 4.71 (SD = 0.49; Range = 3.63-5.88).

Results

Figure S.8.1A displays the proportion of maintenance time participants spent fixating on the location where items had appeared, as a function of value. The proportions were generally low across all value levels. Figure S.8.1B shows the difference in the proportion of maintenance time between value conditions. A one-way repeated measures ANOVA was conducted with value (three levels: High, Equal, Low) as the independent variable and the proportion of maintanence time as the dependent variable. There was a significant effect of value (*GG-corrected* $F(1.02, 34.64) = 20.43$, $MSE < 0.01$, $p < .001$, $\eta_p^2 = .38$; $BF_{10} > 10{,}000$). Participants spent a greater proportion of the maintenance phase fixating on the location where high value items (M = 0.11, SE = 0.01) had appeared, relative to the equal (M = 0.06, SE < 0.01; $p < .001$; d = 0.76; BF_{10} = 297.93) and low value items (M = 0.05, SE < 0.01; $p < .001$, d = 0.77; BF_{10} = 382.92). Participants also spent longer fixating on the location where equal value items had appeared relative to low value items ($p < .001$; d = 0.73; BF_{10} = 185.88).

Figure S.8.1

The Proportion of the Maintenance Phase Participants Spent Fixating on Items in Experiment 1



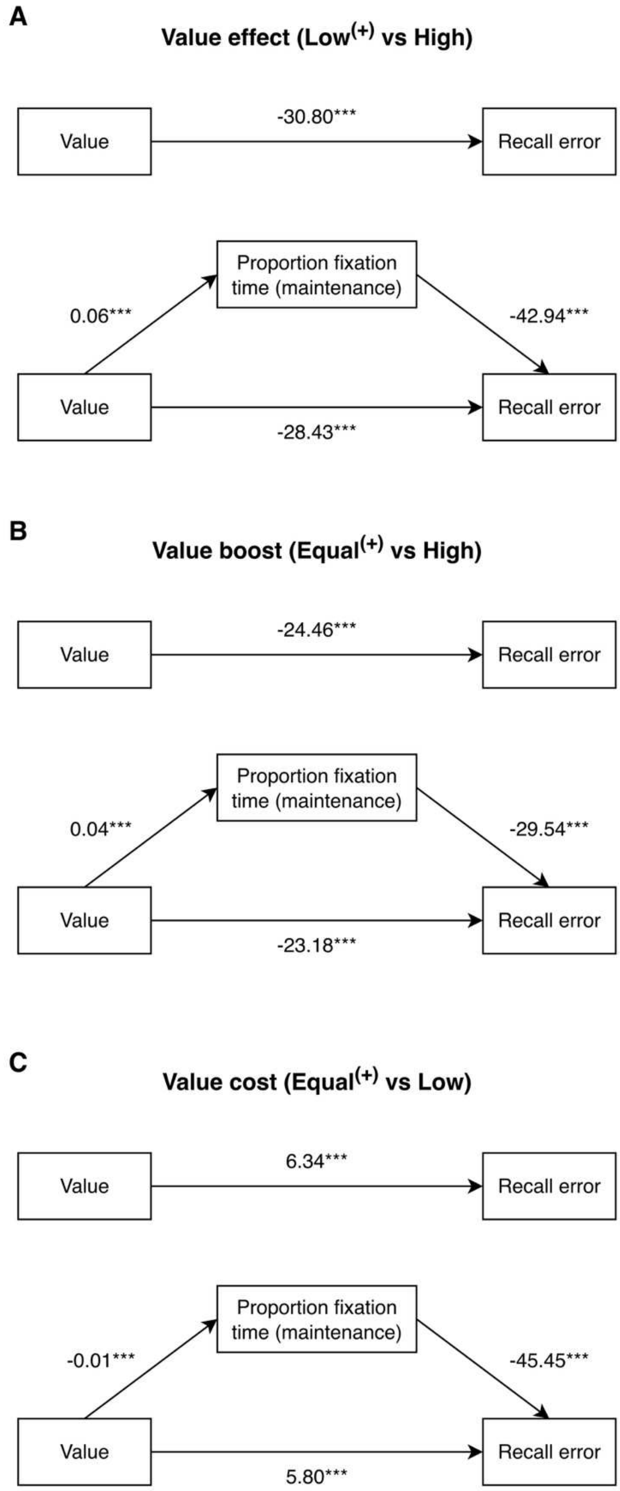
Note. Panel A displays mean proportion of the maintenance time participants spent fixating on an item as a function of value. Panel B displays the mean difference in maintenance time between the value conditions. The dotted line at 0 reflects no difference. Values above 0 indicate that participants fixated on the first condition listed for a higher proportion of the encoding phase relative to the second condition listed. Meanwhile, values below 0 indicate that participants fixated on the first condition listed for a lower proportion of the encoding phase relative to the second level listed. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error of the mean.

Mediation analysis then examined whether the proportion of time participants fixated on the location where the tested item had appeared mediated the relationship between value and recall error. As in the main analyses, separate mediation models were conducted for the overall value effect (low vs high value), the value boost (equal vs high value), and the value cost (equal vs low value). The outcomes are presented in Figure S.8.2. The effect of value on recall error was partially mediated by the proportion of time participants fixated on the location where the tested item appeared during maintenance in all models. The indirect effects were –2.38 for the overall value effect (low vs high; 95% CI: -3.02 - -1.80; $p < .001$), -1.29 for the value boost (equal vs high; 95% CI: -1.90--0.78; $p < .001$), and 0.55 for the value cost (equal vs low; 95% CI: 0.22-0.89; $p < .001$). The proportion of the effect mediated was 0.08 (95% CI: 0.06-0.10; $p < .001$) for the overall value effect (low vs high), 0.05 (95%

CI: 0.03-0.08; $p < .001$) for the value boost (equal vs high), and 0.09 (95% CI: 0.03-0.19; $p < .001$) for the value cost (equal vs low). These outcomes indicate the effect of value is partially mediated by the proportion of time participants spend fixating on the location of an item during maintenance. However, the mediation effect was smaller than that observed when examining the equivalent measure during encoding (see main body).

Figure S.8.2

Outcomes from the Mediation Analysis in Experiment 1 Examining Maintenance Fixation Duration as a Mediator



Note. Panel A reflects the overall value effect (low vs high value), Panel B reflects the value boost (equal vs high value), and Panel C reflects the value cost (equal vs low value). In all panels, the upper diagram reflects the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID.

[+] denotes reference category. *** $p < .001$, ** $p < .01$, * $p < .05$.

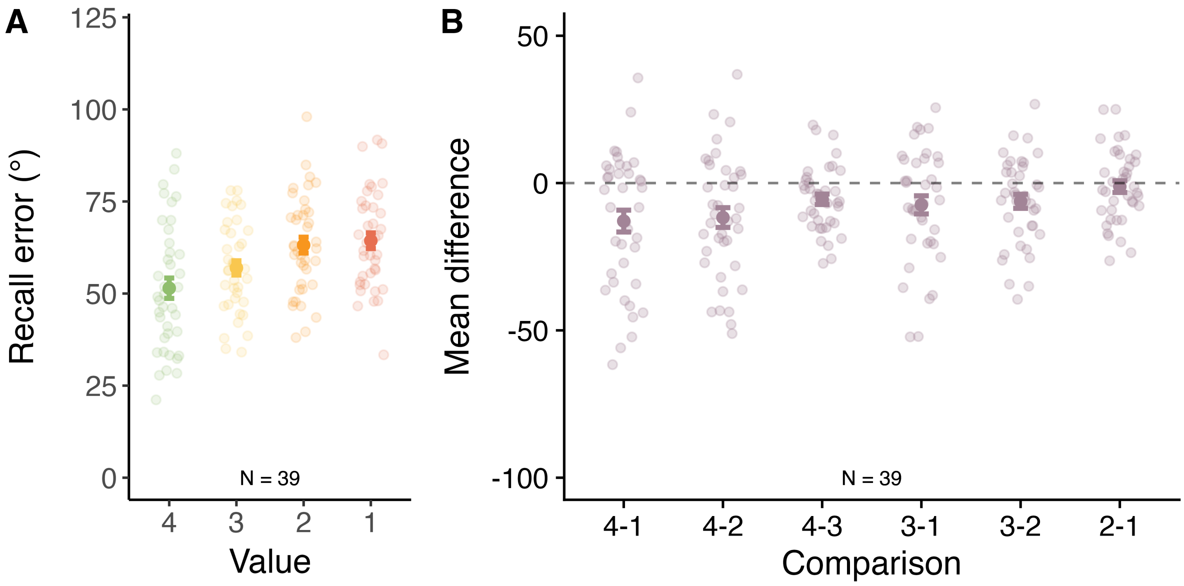
9. Experiment 2: Behavioural outcomes for all participants

This analysis includes all participants with usable behavioural data (*N* = 39). This includes one additional participant excluded from the analysis in the main body due to poor eye-tracking. Mean recall error is displayed in Figure S.9.1.A, as a function of value. Figure S.9.1.B displays the difference in recall error between the value conditions. A one-way repeated measures ANOVA was conducted to investigate the effect of value on recall error. There was a significant effect of value (*GG-corrected* $F(1.68, 63.87) = 9.04$, $MSE = 274.89$, $p < .001$, $\eta_p^2 = .19$; $BF_{10} = 1401.48$). Post-hoc tests revealed that recall error for items worth 4 points ($M = 51.4$, $SE = 2.77$) was significantly lower than items worth 1 point ($M = 64.4$, $SE = 2.13$, $p = .007$, $d = -0.56$; $BF_{10} = 24.71$), 2 points ($M = 63.2$, $SE = 2.16$; $p = .007$, $d = -0.56$; $BF_{10} = 25.47$) and 3 points ($M = 57.0$, $SE = 1.94$; $p = .014$, $d = -0.50$; $BF_{10} = 9.91$). Furthermore, recall error was marginally significantly lower for items worth 3 points relative to those worth 1 point ($p = .043$, $d = -0.38$; $BF_{10} = 2.16$) and 2 points ($p = .042$, $d = -0.41$; $BF_{10} = 3.11$). As in the main analyses, there was no significant difference between items worth 2 and 1 points ($p = .551$, $d = -0.10$), with BF evidence against an effect ($BF_{10} = 0.20$).

Thus, the conclusions were identical to those reported in the main body (except that recall error was marginally significantly lower for items worth 3 points than those worth 1 and 2 points in the current analyses, but not in the analysis reported in the main body).

Figure S.9.1

Task Performance in Experiment 2 for all Participants With Usable Task Data (N = 40)



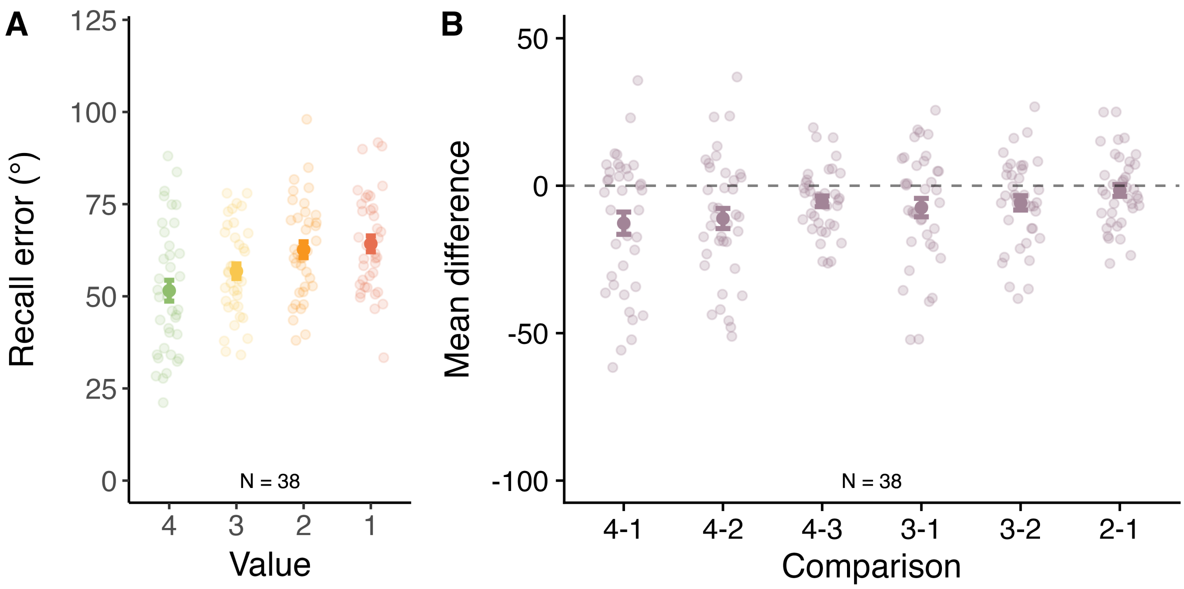
Note. Panel A displays mean recall error, as a function of value (higher recall error reflects worse performance). Panel B displays the mean difference between the value conditions, calculated as the first condition listed minus the second (e.g. the figure for 4 vs 1 for a given participant was calculated by subtracting their mean recall error in the "1" point value condition from their mean recall error in the "4" point value condition). Across panels, bold points reflect mean across participants, whilst error bars denote standard error. Lighter circles in the background indicate individual participants' mean.

10. Experiment 2: Behavioural outcomes for only trials with fixations recorded for participants analysed in main body (N = 38).

This analysis reports behavioural outcomes for participants reported in the main body, but focusing only on trials with usable eye-tracking data (N = 38). Mean recall error as a function of value is displayed in Figure S.10.1.A, whilst the mean difference in recall error between conditions is displayed in Figure S.10.1.B. A one-way repeated measures ANOVA revealed a significant effect of value (*GG-corrected* $F(1.67, 61.88) = 8.25$, $MSE = 278.64$, $p < .001$, $\eta_p^2 = .18$; $BF_{10} = 571.48$). Post-hoc tests revealed that recall error for items worth 4 points (M = 51.5, SE = 2.83) was significantly lower than items worth 1 point (M = 64.2, SE = 2.16, p = .011, d = -0.55; BF_{10} = 18.28), 2 points (M = 62.6, SE = 2.16; p = .012, d = -0.53; BF_{10} = 13.99) and 3 points (M = 56.8, SE = 1.99; p = .025, d = -0.47; BF_{10} = 6.32). None of the other comparisons were significant following correction ($p \geq .065$; d = -0.13 - -0.39; $BF_{10} \leq 2.15$). Further, as in the main analyses, the BF analysis revealed evidence against a difference between items worth 1 and 2 points (BF_{10} = 0.23). Therefore, the conclusions from this analysis did not differ from those presented in the main body.

Figure S.10.1

Task Performance in Experiment 1 for only Trials With Eye-Tracking Fixations Recorded During the Encoding Phase (N = 39)



Note. Panel A displays mean recall error, as a function of value (higher recall error reflects worse performance). Panel B displays the mean difference between the value conditions, calculated as the first condition listed minus the second (e.g. the figure for 4 vs 1 for a given participant was calculated by subtracting their mean recall error in the “1” point value condition from their mean recall error in the “4” point value condition). Across panels, bold points reflect mean across participants, whilst error bars denote standard error. Lighter circles in the background indicate individual participants’ mean.

11. Experiment 2 response time analyses (*N* = 38)

As in Experiment 1, we conducted analyses of the response times to rule out a speed-accuracy trade off. This was conducted using the data for all trials in participants included in the final sample.

Pre-processing

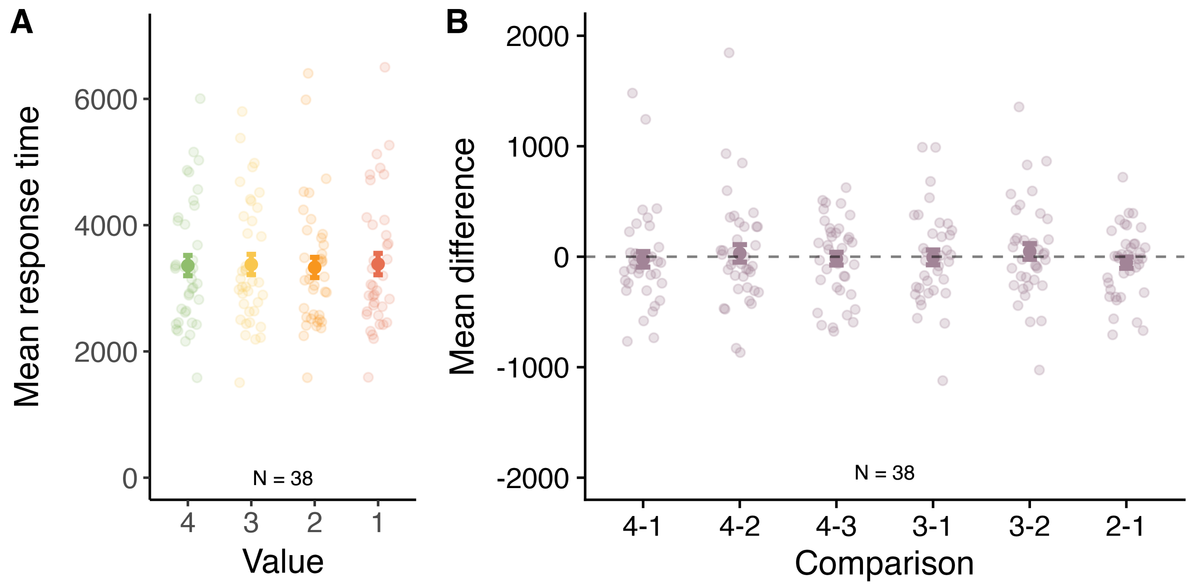
The same pre-processing steps were applied as in Experiment 1: we first excluded trials where the response time was over 20,000ms (20 seconds; 10 trials excluded) and then excluded trials that were 2.5 SDs above or below the mean for each value condition for each participant.

Results

Mean response times as a function of value are displayed in Figure S.11.1.A, whilst Figure S.11.1.B displays the difference in response times between the three value conditions. A one-way repeated measures ANOVA was conducted to investigate the effect of value (three levels: High, Equal, Low) on response times. There was no significant effect of value (GG-corrected, $F(2.39, 88.38) = 0.26$, $MSE = 106723.55$, $p = .809$, $\eta_p^2 = .01$; $BF_{10} = 0.05$, $BF_{01} = 21.39$).

Figure S.11.1

Response Times in Experiment 2



Note. Panel A displays mean response times, as a function of value. Panel B displays the mean difference between the three value conditions. The dotted line at 0 reflects no difference. Values above 0 indicate slower response times in the first condition listed, whilst values below 0 indicate faster response times in the first condition listed. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error (SE) of the mean.

12. Experiment 2: Analysis of the proportion of fixations and length of individual fixations

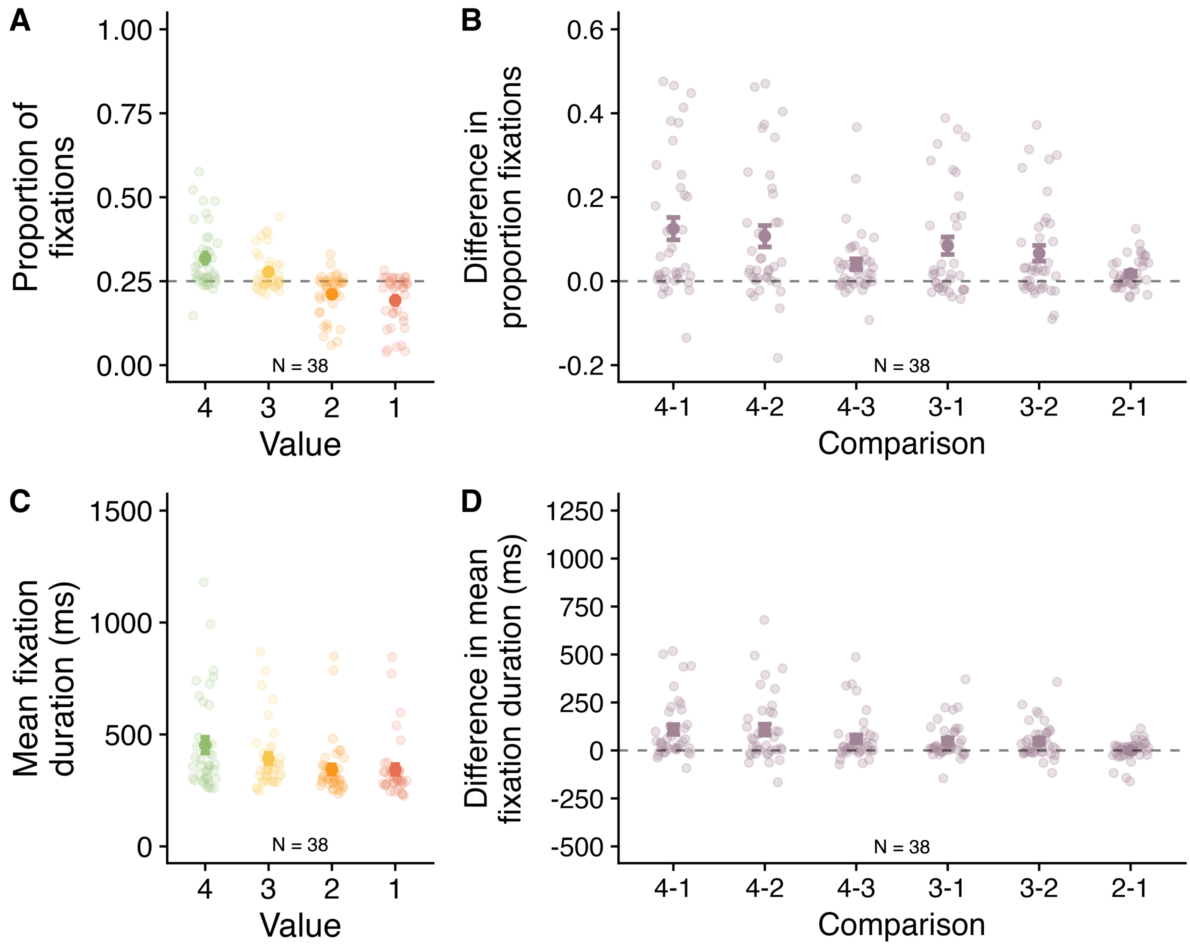
Proportion of fixations

We first analyzed the proportion of fixations (a measure which ignores the length of fixations), which was calculated in the same manner as in Experiment 1. Given that all items differed in their value in this experiment, separate proportions were computed for each of the four value levels. The mean proportion of fixations during encoding is displayed in Figure S.12.1.A, as a function of item value. The difference between value condition pairs is displayed in Figure S.12.1.B. Firstly, a one-way ANOVA was conducted to examine whether the proportion of fixations towards items differed as a function of value (four levels [within-subject]: 1-4 points). This revealed a significant effect of value (GG-corrected $F(1.29, 47.71) = 17.21$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = .32$, $BF_{10} > 10{,}000$). Post-hoc tests revealed that the proportion of fixations during the encoding phase was higher for items associated with four points ($M = 0.32$, $SE = 0.02$) relative to those associated with one point ($M = 0.19$, $SE = 0.01$; $p < .001$; $d = 0.76$; $BF_{10} = 583.52$), two points ($M = 0.21$, $SE = 0.01$; $p = .001$; $d = 0.68$; $BF_{10} = 154.17$), and three points ($M = 0.28$, $SE = 0.01$, $p = .006$, $d = 0.51$; $BF_{10} = 11.61$). Moreover, the proportion of fixations towards items associated with three points was greater than those associated with one point ($p = .001$, $d = 0.65$; $BF_{10} = 102.21$) and two points ($p = .003$, $d = 0.58$; $BF_{10} = 29.01$). Finally, the proportion of fixations towards items worth two points was greater than those worth one point ($p = .006$, $d = 0.49$; $BF_{10} = 7.82$).

A series of four one-sample t-tests (corrected using Bonferroni-Holm) were then performed comparing the proportion in each value condition to the value expected given an equal distribution of attention to all items (0.25). The proportion of fixations on the items worth four points ($t(37) = 4.44$, $p < .001$, $d = 0.72$; $BF_{10} = 305.47$) and three points ($t(37) = 2.90$, $p = .006$, $d = 0.47$; $BF_{10} = 6.18$), were significantly greater than 0.25. The opposite pattern emerged for items worth two points ($t(37) = -3.66$, $p = .002$, $d = -0.59$; $BF_{10} = 38.74$) and one point ($t(37) = -4.64$, $p < .001$, $d = -0.75$; $BF_{10} = 532.96$), with proportions significantly less than 0.25.

Figure S.12.1

Proportion of Fixations and Length of Individual Fixations in Experiment 2



Note. Panel A: Mean proportion of fixations, as a function of value. The dotted horizontal line at 0.25 reflects the proportion expected if fixations were equally distributed across the four-item display. Panel B: Mean difference in the proportion of fixations for each value pair comparison. The dotted horizontal line at 0 reflects no difference. Values above 0 indicate that the proportion of fixations were higher for the first condition listed relative to the second. Meanwhile, values below 0 indicate that the proportion of fixations were lower for the first condition listed relative to the second. Panel C: Mean duration per fixation, as a function of value. Panel D: Mean difference in the duration per fixation for each value pair comparison. dotted horizontal line at 0 reflects no difference. Values above 0 indicate that participants' individual fixations were longer for the first condition listed relative to the second. Meanwhile, values below 0 indicate that participants' individual fixations were shorter for the first condition listed relative to the second. For all panels, error bars denote standard error and lighter dots in the background reflect means for individual participants.

Length of individual fixations

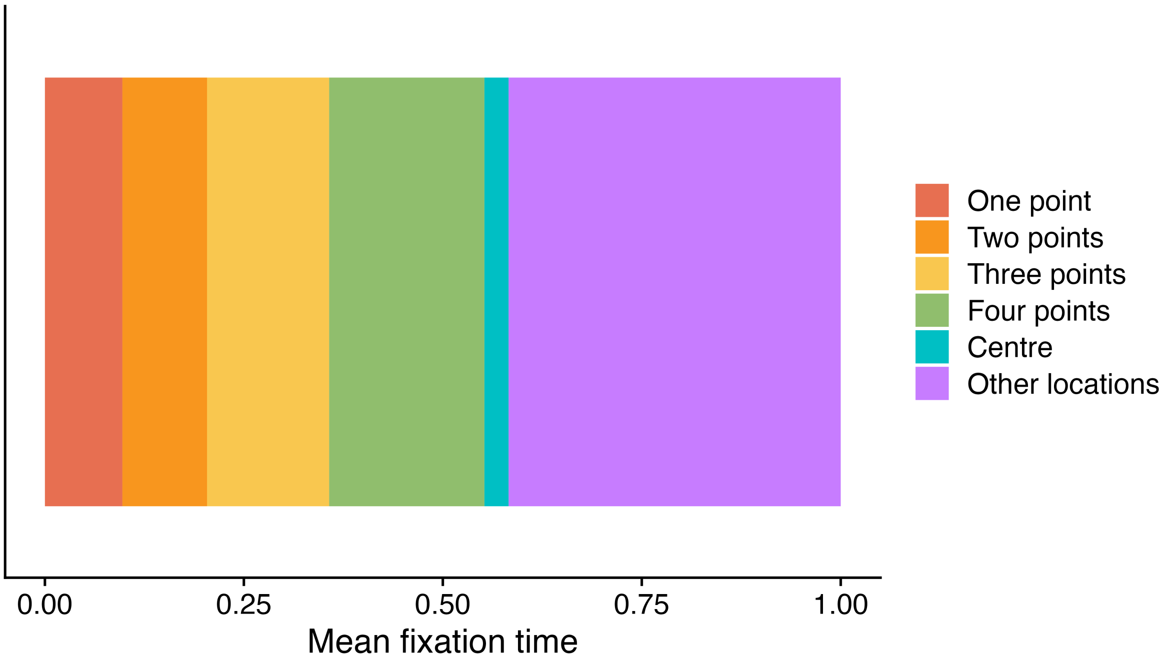
We next examined whether the mean length of individual fixations differed as a function of value. This analysis was comparable to the analysis performed in Experiment 1. The mean duration of individual fixations is displayed in Figure S.12.1.C, whilst the difference in fixation length between value condition pairs is shown in Figure S.12.1.D. A one-way repeated measures ANOVA (four levels: 1-4 points) found a significant effect of value (GG-corrected $F(1.49, 55.14) = 13.37$, $MSE = 15234.10$, $p < .001$, $\eta_p^2 = .27$, $BF_{10} > 10{,}000$). Post-hoc tests revealed that the mean fixation length was longer for items associated with four points ($M = 452$, $SE = 35.1$) than for items associated with one point ($M = 343$, $SE = 21.9$, $p = .001$, $d = 0.69$; $BF_{10} = 187.54$), two points ($M = 344$, $SE = 20.6$, $p = .003$, $d = 0.62$; $BF_{10} = 57.06$), and three points ($M = 392$, $SE = 23.8$, $p = .011$, $d = 0.49$; $BF_{10} = 8.55$). Moreover, fixations towards items associated with three points were longer than those associated with one point ($p = .011$, $d = 0.52$; $BF_{10} = 12.15$) and two points ($p = .011$; $d = 0.52$; $BF_{10} = 12.99$). There was no significant difference in the length of fixations to items associated with one and two points ($p = .963$; $d = 0.01$), with BF evidence in favor of no effect ($BF_{10} = 0.17$).

13. Fixations including screen centre as an area of interest (AOI; Experiment 2)

Figure S.13.1 shows the mean proportion of the encoding phase participants spent fixating on the spatial location of items worth 1 point, 2 points, 3 points, 4 points, the screen centre[10], and other areas of the screen.

Figure S.13.1

Mean Proportion of the Encoding Phase Participants Spent Fixating on the Items (as a Function of Value), the Screen Centre, and Other Areas of the Screen



Note. Unlike in Experiment 1, data for 1 point, 2 points, 3 points, 4 points all reflect fixations to only one item per trial (as all items presented in each trial in this experiment had a different value).

[10] The area of interest (AOI) was 350x350 pixels around the screen centre.

14. Encoding fixation duration as a mediator (further analysis; Experiment 2)

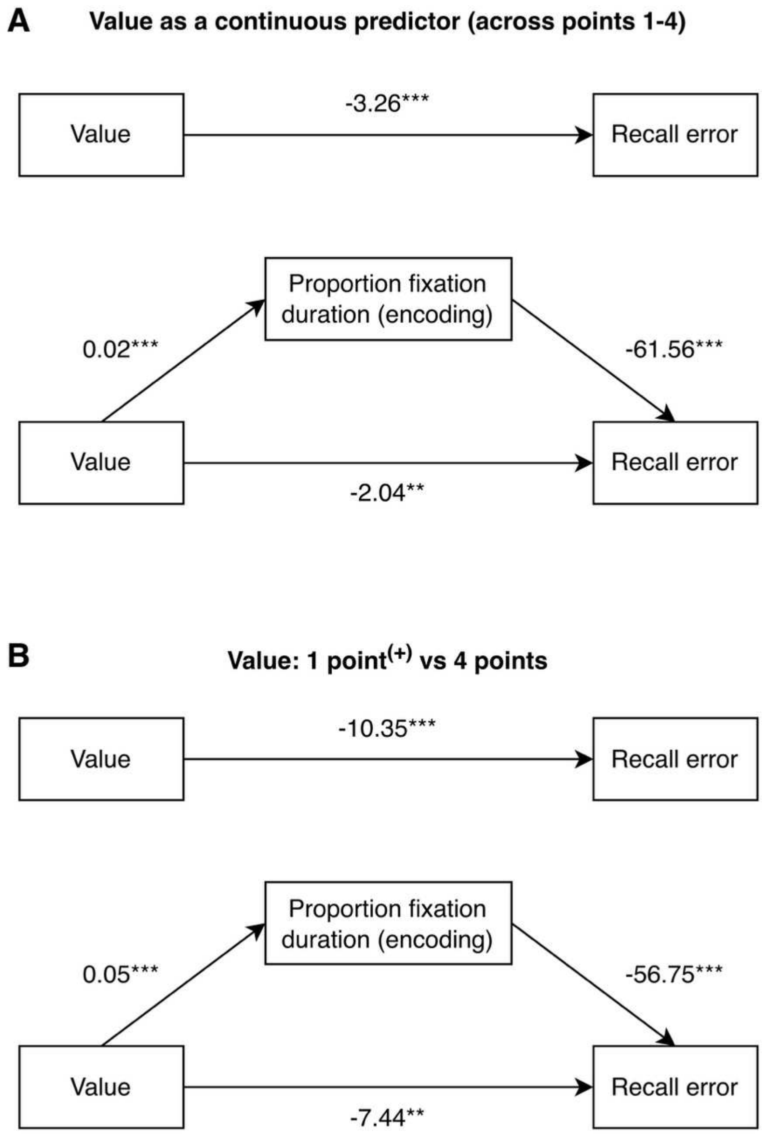
A pair of mediation models were conducted to investigate whether encoding fixation duration still partially mediated the effect of value on recall error after trials were excluded whether the first fixation was on the tested item. This investigates whether encoding fixation duration mediates the relationship after accounting for a primacy effect. The first model included value as a continuous variable (1-4). The second model compared values 1 and 4. The outcomes are presented in Figure S.14.1.

For the model with value as a continuous variable, the indirect effect was -1.23 (95% CI: --1.56--0.93; $p < .001$), and the proportion mediated was 0.38 (95% CI: 0.25-0.65; $p < .001$). For the model comparing values 1 vs 4 only, the indirect effect was -0.99 (95% CI: -1.39--0.64; $p < .001$), and the proportion mediated was 0.28 (95% CI: 0.17-0.52; $p < .001$).

These outcomes indicate the effect of value on recall error is partially mediated by the proportion of time participants spend fixating on an item during encoding, even after when removing trials where the first item fixated on was the tested item.

Figure S.14.1

Outcomes from the Mediation Analysis in Experiment 2 Examining Encoding Fixation Duration as a Mediator, After Excluding Trials Where the First Fixation was on the Tested Item



Note. Panel A: Outcomes from the mediation model, when considering all four value conditions (values 1-4), entered into the model as a continuous predictor. Panel B: Outcomes from a mediation model, comparing "1" and "4" point values (comparable to the "Low" vs "High" mediation analyses in Experiment 1). In this latter model, the "1" point condition was entered as the reference category (denoted in the figure by (+)). In both panels, the upper diagram displays the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. *** $p < .001$, ** $p < .01$, * $p < .05$.

15. Experiment 2: Maintenance fixation duration

As in Experiment 1, analysis was performed to investigate whether the proportion of maintenance time participants spent fixating on locations in which items had appeared differs as a function of value; and whether this factor mediates the relationship between value and recall error.

Data processing

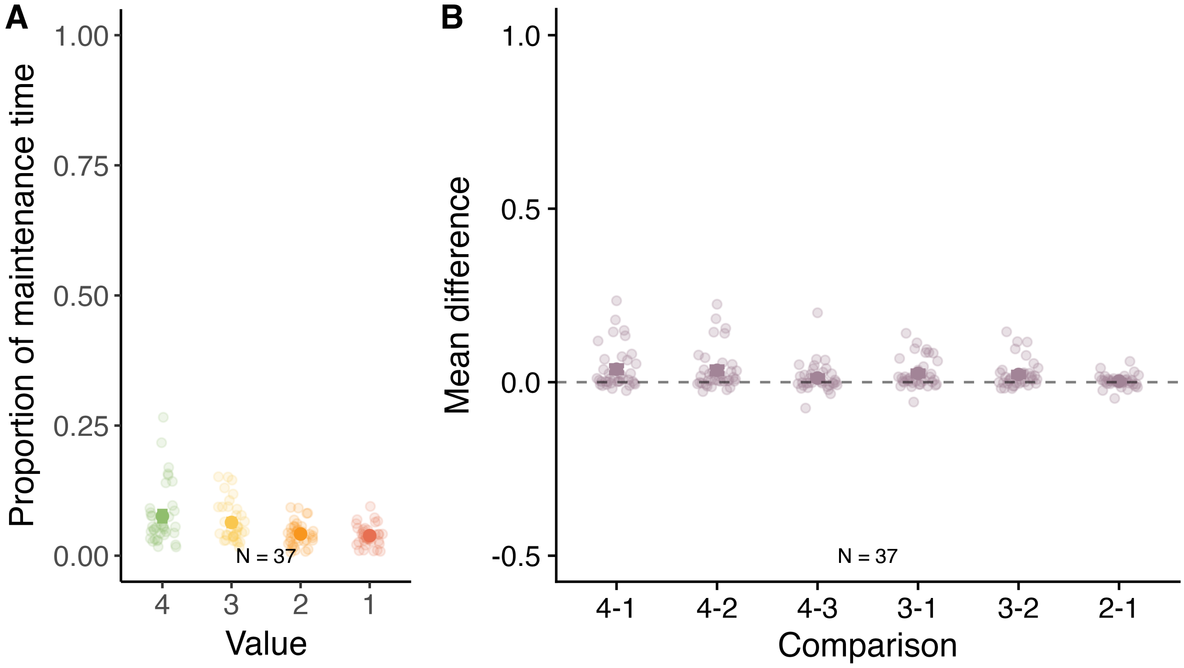
In addition to the one participant excluded in the main analysis (see *Participant* section of Experiment 1), one further participant was excluded as the number of trials with no fixations during maintenance was over 2.5 standard deviations above the mean. The analysis was therefore performed on data for 37 participants. The percentage of trials with no recorded fixations was low in this dataset (M = 0.84, SD = 1.66). Trials with no recorded fixations were excluded. Following this, the mean number of fixations across participants in the maintenance phase was 4.47 (SD = 0.67; Range = 2.88-5.62).

Results

Figure S.15.1A displays the proportion of maintenance time participants spent fixating on the location where items had appeared, as a function of value. Figure S.15.1B shows the difference in the proportion of maintenance time between value conditions. A one-way repeated measures ANOVA was conducted with value (three levels: High, Equal, Low) as the independent variable and the proportion of maintenance time as the dependent variable. There was a significant effect of value (*GG-corrected* $F(1.56, 56.28) = 11.06$, $MSE < 0.01$, $p < .001$, $\eta_p^2 = .24$; $BF_{10} > 10{,}000$). Participants spent a significantly higher proportion of the maintenance phase fixating on the location where an item worth 4 points had appeared (M = .08, SE = .01), relative to those worth 1 point (M = .04, SE < .01; p = .005; d = 0.60; BF_{10} = 38.48) and 2 points (M = .04, SE < .01; p = .006; d = 0.57; BF_{10} = 23.72). Participants also spent a greater proportion of time fixating on the location where an item worth 3 points (M = .06, SE = .01) had appeared relative to those worth 1 point (p = .005, d = 0.59; BF_{10} = 32.00) or 2 points (p = .006, d = 0.56; BF_{10} = 20.28). No significant differences emerged between items associated with 4 and 3 points (p = .175, d = 0.29; BF_{10} = 0.71, BF_{01} = 1.41), or 2 and 1 points (p = .293, d = 0.18; BF_{10} = 0.30; BF_{01} = 3.34).

Figure S.15.1

The Proportion of the Maintenance Phase Participants Spent Fixating on Items in Experiment 2

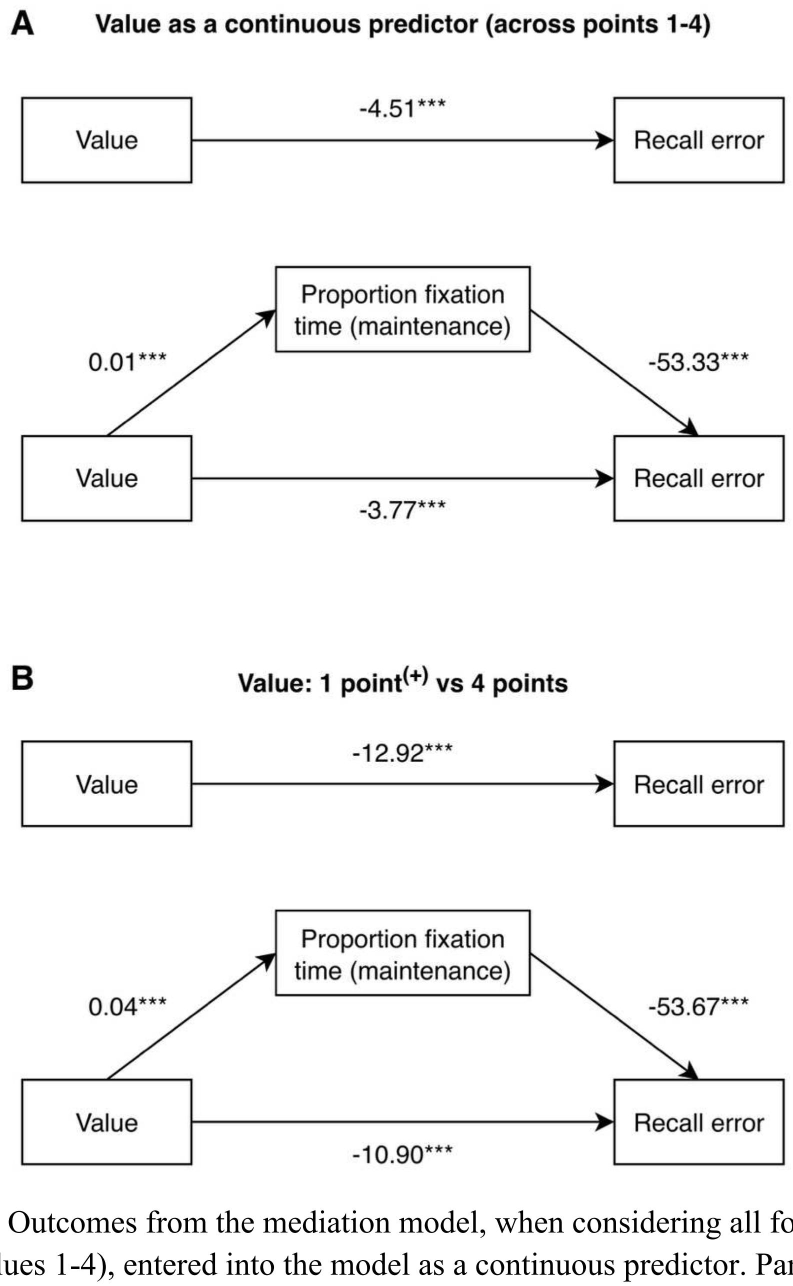


Note. Panel A displays mean proportion of the maintenance time participants spent fixating on an item as a function of value. Panel B displays the mean difference in maintenance time between the value conditions. The dotted line at 0 reflects no difference. Values above 0 indicate that participants fixated on the first condition listed for a higher proportion of the encoding phase relative to the second condition listed. Meanwhile, values below 0 indicate that participants fixated on the first condition listed for a lower proportion of the encoding phase relative to the second level listed. Across panels, bold points reflect the mean across participants, while lighter circles in the background show each participant's mean. Error bars show +/- 1 standard error of the mean.

Mediation analysis then examined whether the proportion of time participants fixated on the location where the tested item had appeared mediated the relationship between value and recall error. The same models were conducted as in the main analyses described in the main body (i.e. treating value as continuous, and examining the lowest vs highest value levels (1 and 4 points)). The outcomes are displayed in Figure S.15.2. For the model treating value as continuous, the indirect effect was -0.74 (95% CI: -0.95--0.56; $p < .001$), and the proportion of the effect mediated was 0.16 (95% CI: 0.12-0.23; $p < .001$). For the model comparing the lowest and highest value levels, the indirect effect was -2.03 (95% CI: --2.76--1.38; $p < .001$), and the proportion of the effect mediated was 0.16 (95% CI: 0.10-0.23, $p < .001$). Therefore, the effect of value on recall error was partially mediated by the proportion of maintenance time participants fixated on the location where the tested item appeared. However, as in Experiment 1, the proportion of the effect explained by fixation time during maintenance was considerably smaller than that explained by the fixation time during encoding.

Figure S.15.2.

Outcomes from the Mediation Analysis in Experiment 2, Investigating Proportion of Fixation Time during Maintenance as the Mediator.



Note. Panel A: Outcomes from the mediation model, when considering all four value conditions (values 1-4), entered into the model as a continuous predictor. Panel B: Outcomes from a mediation model, comparing "1" and "4" point values (comparable to the "Low" vs "High" mediation analyses in Experiment 1). In this latter model, the "1" point condition was entered as the reference category (denoted in the figure by (+)). In both panels, the upper diagram displays the total effect of value on recall error. The lower diagram shows the mediation analysis. All models included a random intercept of participant ID. + denotes reference category. *** $p < .001$, ** $p < .01$, * $p < .05$.