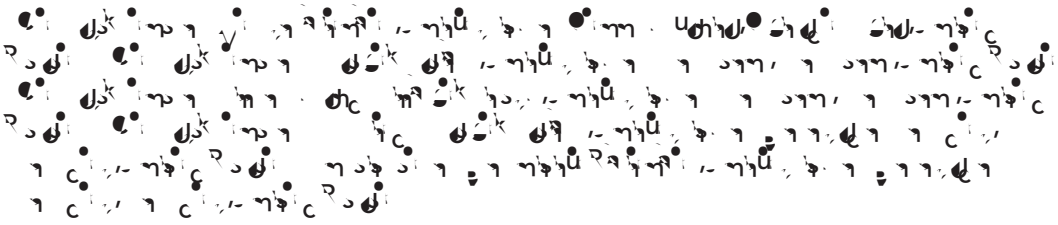
Strategy-dependent effects of working-memory limitations on human perceptual decision-making

Kyra Schapiro[1]*, Krešimir Josić[2,3], Zachary P Kilpatrick[4,5], Joshua I Gold[1]



Abstract Deliberative decisions based on an accumulation of evidence over time depend on working memory, and working memory has limitations, but how these limitations affect deliberative decision-making is not understood. We used human psychophysics to assess the impact of working-memory limitations on the fidelity of a continuous decision variable. Participants decided the average location of multiple visual targets. This computed, continuous decision variable degraded with time and capacity in a manner that depended critically on the strategy used to form the decision variable. This dependence reflected whether the decision variable was computed either: (1) immediately upon observing the evidence, and thus stored as a single value in memory; or (2) at the time of the report, and thus stored as multiple values in memory. These results provide important constraints on how the brain computes and maintains temporally dynamic decision variables.

Editor's evaluation

This paper employs sophisticated modeling of human behavior in well-controlled tasks to study how limitations of working memory constrain decision-making. Because both are key cognitive processes, that have so far largely been studied in isolation, the paper will be of broad interest to neuroscientists and psychologists. The observed working memory limitations support previous findings and extend them in critical ways.

Introduction

Many perceptual, memory-based, and reward-based decisions depend on an accumulation of evidence over time (***Brody and Hanks, 2016***; ***Gold and Shadlen, 2007***; ***Ratcliff et al., 2016***; ***Shadlen and Shohamy, 2016***; ***Summerfield and Tsetsos, 2012***). This dynamic process, which can operate on timescales ranging from tens to hundreds of milliseconds for many perceptual decisions to seconds or longer for reward-based and other decisions (***Bernacchia et al., 2011***; ***Gold and Stocker, 2017***), requires working memory to maintain representations of new, incoming evidence and/or the aggregated, updating decision variable. Working memory is constrained by capacity and temporal limita-

For spatial working-memory tasks, the precision of working memory for perceived spatial locations is often well described by diffusion dynamics (***Compte et al., 2000***; ***Kilpatrick, 2018***; ***Kilpatrick et al., 2013***; ***Laing and Chow, 2001***) that are commonly implemented in 'bump-attractor' models of working memory (***Compte et al., 2000***; ***Constantinidis et al., 2018***; ***Laing and Chow, 2001***; ***Riley and Constantinidis, 2016***; ***Wei et al., 2012***; ***Wimmer et al., 2014***). Our analyses built on this framework by examining memory diffusion dynamics for the different task conditions and potential decision strategies. For the conditions we tested, most participants' behavior was well fit by one of two distinct strategies, each with its own constraints on decision performance based on different working-memory demands. The first strategy was to compute the decision variable (mean disk angle) immediately upon observing the evidence (individual disk angles), and then store that value in working memory in a manner that, like for the memory of a single perceived angle, could be modeled as a single particle with a particular diffusion constant (Average-then-Diffuse model; AtD). The second strategy was to maintain representations of all disk locations in working memory, modeled as separate diffusing particles, and then to combine them into a decision variable only at the time of the decision (Diffuse-then-Average model; DtA). Such a strategy results in an effective diffusion constant for the average that is inversely related to the number of items. Our results show that like perceived locations, memory for computed mean locations degraded with increased set size (of relevant information), and delay between presentation and report. However, the degree of degradation depended on the strategy used to compute the decision variables, implying that multip utinct lb0.5ai. On P

we measured the error between reported and probed angles as a proxy for working-memory representations and inferred rates of memory degradation (diffusion constants) from the increase in variance of these errors over time within a framework of diffusing-particle models. Below we first describe the model framework, detailing its key assumptions and predictions. We next describe results from Simultaneous conditions, in which all items were presented simultaneously at the beginning of each trial, which demonstrate how capacity and temporal constraints on working memory relate to the accuracy of computed decision variables. We then describe results from Sequential conditions, in which one item was presented after the others in each trial, which demonstrate how capacity and temporal constraints affect the process of evidence integration over time.

Diffusing-particle framework and predictions

Within our diffusing-particle framework, the memory of an item is represented by the location of a diffusing particle. This representation allows us to quantify the corruption (i.e., reduced precision) of the memory by two distinct sources of noise. The first is described by a static, additive term (η_1) that encompasses all potential one-time noise sources within a trial including noise associated with the sensory encoding and the motor response. The second is the dynamic degradation of memory precision over time that is modeled as the diffusion of the particle (***Figure 2a***). This diffusion corresponds to an increase in variability over time that is linear, with a slope equal to the diffusion constant (σ_1^2; ***Figure 2b***). Consistent with past modeling studies (***Bays et al., 2009***; ***Brady and Alvarez, 2015***; ***Koyluoglu et al., 2017***; ***Wei et al., 2012***), we accounted for the decrease in working-memory fidelity with item load by incorporating item noise (N

of averaging, and AtD produces a lower σ_{MN}^2 and less variable responses than DtA. A summary of all framework variables can be found in ***Table 1***.

To summarize, our two models describe two different possible ways for decision-relevant information to be stored in working memory prior to executing a decision. The different storage strategies result in different patterns of memory degradation, corresponding to trial-to-trial variability (imprecision) of decision reports that increase as a function of the length of the within-trial delay period. For

Table 1. Descriptions of all framework and model parameters.
Fit parameters are shown on the top. Derived parameters used in other analyses and descriptions are shown on the bottom. Variations used to model Sequential conditions are shown to the right. *p*

the AtD model, the individual pieces of information are immediately combined into a single decision variable that is then stored in memory. Thus, the rate of degradation of an estimated average is identical to the rate of degradation of a single item. In contrast, for the DtA model, all of the relevant pieces of information are stored in memory and then combined only at the time of indicating the decision. Thus, the rate of degradation of an estimated average is inversely proportional to the rate of degradation of each item held in memory. We used fits of these models to performance data from individual participants to distinguish different patterns of memory degradation and therefore different storage strategies.

Simultaneous condition behavior

When all disks were presented simultaneously, performance was consistent with several key predictions of the particle model. Specifically, the difference in reports of Perceived spatial angles and the true probed location (i.e., the response error) tended to be unbiased, in that the mean error across participants was not reliably different from 0 (***Figure 3a***, full distributions in ***Figure 3—figure supplement 1***, individual participant mean errors in ***Figure 3—figure supplement 2***). However, the variance of these errors increased roughly linearly over time (***Figure 3c***), like the location of a diffusing particle or bump attractor (***Compte et al., 2000***; ***Kilpatrick, 2018***; ***Kilpatrick et al., 2013***; ***Laing and Chow, 2001***). This error variance depended systematically on set size (***Figure 3c***). However, the change in error variance over time (slope of variance increase) did not depend on set size (ANOVA, significant effect of set size, $F(2,32)=83.87$, $p=1.88e{-}13$, and delay, $F(2,32)=29.55$, $p=5.37e{-}08$, but no significant interaction between set size and delay, $F(4,64)=1.36$, $p=0.256$). Errors in reports of Computed (i.e., inferred mean) spatial angles relative to true mean angles showed similar trends, albeit with a much weaker dependence on the number of items. Specifically, Computed angle reports were also unbiased (mean error from the true value was not reliably different from 0; ***Figure 3b***, ***Figure 3—figure supplements 1 and 3***) but degraded (became more variable) with a roughly linear increase in variance over time (***Figure 3d***). Error variance in the report of the Computed average was higher at higher set sizes (set size 5 had higher variances), but the rate of degradation in accuracy did not depend on set size (ANOVA, significant effect of set size, $F(2,32)=13.53$, $p=5.515e{-}5$, and delay, $F(2,32)=130.79$, $p=4.441e{-}16$, but not their interaction, $F(4,64)=0.538$, $p=0.708$).

summarized in ***Figure 4c and f***, thus support the idea that for most participants, their behavior was well captured by their better-fitting model.

Summaries of the predicted report-error variances by the AtD and DtA fits for well-fit participants are shown in ***Figure 5***. Overall, the model predictions qualitatively match participant behavior.
In general, [(Fihehavior was)redicted ry tdif18 (efusons cns tnts ahat fwe)18 (e she msamefor weihe r one
lgure 5.).DtA ehavior was edicted ry tdif18 1 (efusons cns tnts ahat fwe)18 (e slr)18 (eg r or moultiple)]TJ

p=0.029). We also found these results were robust to uncertainty associated with model identifiability (participant-wise identifiability is given in ***Figure 2—figure supplement 1***). Specifically, given different possible distributions of underlying strategy prevalence (proportions), the probability of obtaining the empirically observed distributions of models shown in ***Figure 6a*** for either set size while considering the average model identifiability was peaked near the observed strategy proportions. This result demonstrates that the observed proportions were not likely obtained due misidentification-related chance. These probabilities distributions were also highly non-overlapping, which is consistent with a different prevalence of strategy use at the two different set sizes (***Figure 6b***).

These differences in strategy use did not correlate with the ages of the participants (Pearson correlation, ***Figure 6—figure supplement 1***, p>0.20). These findings suggest that working-memory load might have affected our participants' decision strategies, such that a higher load corresponded to an increased tendency to discard information about individual samples (disk locations) and hold only the relevant computed decision variable in memory.

Sequential condition behavior

For the Sequential condition, we separately analyzed errors for Perceived reports of disks presented at the beginning (Early) or middle (Late) of a trial. Early Perceived reports tended to be relatively unbiased (two-sided *t*-test for H_0: mean error=0, p>0.05; ***Figure 7a***, full distributions in ***Figure 7—figure supplement 1***; individual participant mean errors in ***Figure 7—figure supplement 2a-d***) but became more variable over time in a roughly linear manner (***Figure 7d***), consistent with the predictions of the particle-

eLi

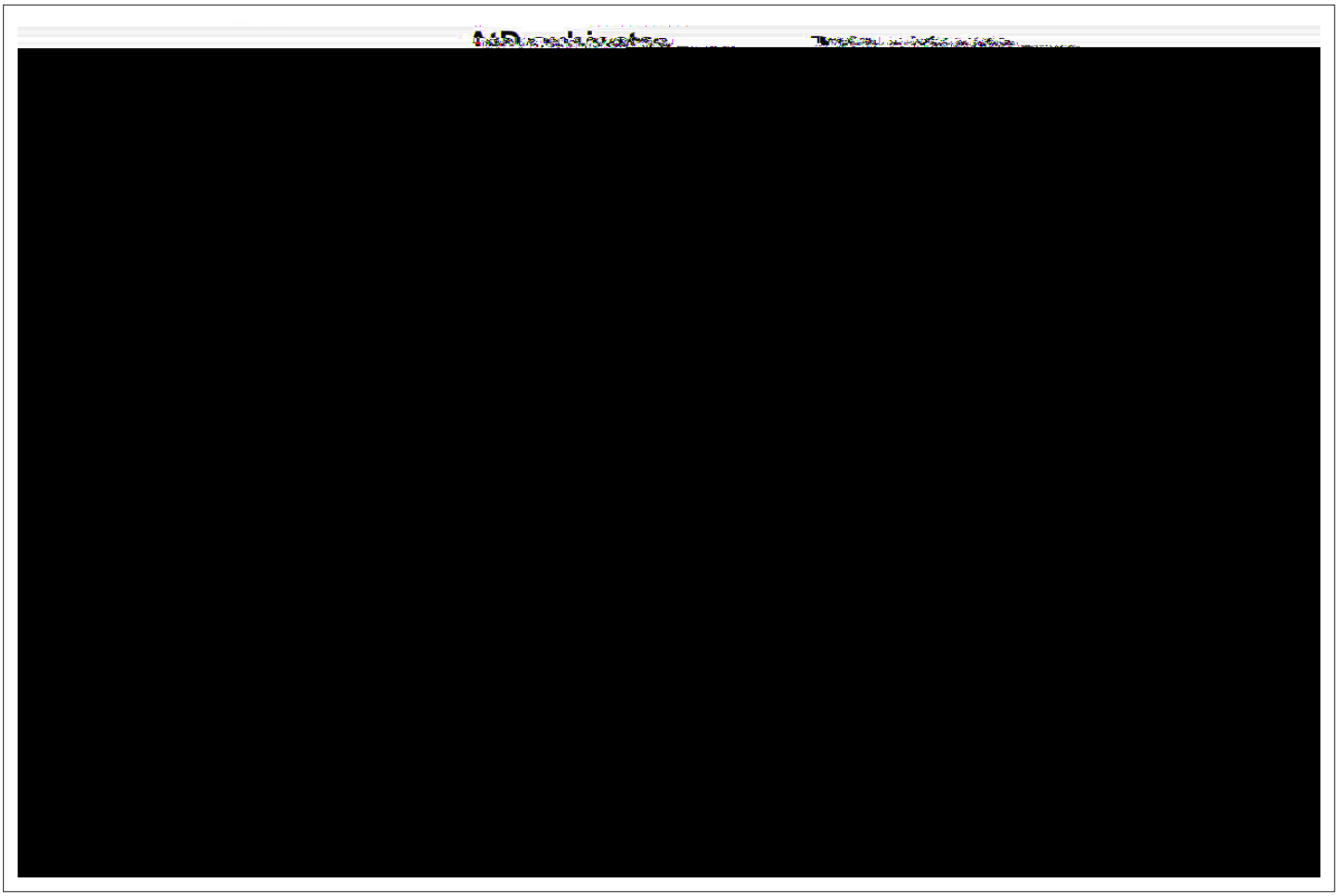


Figure 8. Comparisons of empirical and model-based diffusion constants. In (**a, b, d, e**), the abscissa shows the difference between: (1) empirical estimates of the diffusion constant for a Computed value measured by tting a line to measured variance as a function of delay time for set size 2 (σ_{M2}^2, **a, b**) or 5 (σ_{M5}^2, **d, e**), and (2) the empirical estimates of the diffusion constant for a single Perceived value (σ_1^2) multiplied by the appropriate factor for the set size. The AtD model predicts a difference of 0. The ordinate shows the difference between: (1) the empirical estimate of Computed diffusion constants σ_{M2}^2 or σ_{M5}^2, and (2) the empirical estimates of the diffusion constant of a Computed value based on the DtA hypothesis. The DtA model predicts a difference of 0. Points are data from individual participants, separated by whether they were best t by the AtD (**a, b**) or DtA (**d, e**) model for the given set-size condition. Lines are 95% con dence intervals (CIs) computed by simulating data using the best-t parameters for the given t and repeating the empirical estimate comparison procedure. Close symbols indicate participants who fell within the 95% CI for their best-t model. (**e, f**) Distance of each participant's empirically estimated diffusion constant relationships from those predicted by AtD or DtA (i.e., distances from the =0 and =0 lines, respectively, in (**a, b, d, e**)), for set sizes 2 (**c**) and 5 (**f**). AtD, Average-then-Diffuse; DtA, Diffuse-then-Average.

The online version of this article includes the following gure supplement(s) for gure 8:

Figure supplement 1. Participant-speci c estimates of *A* from the Sequential condition for set sizes 2 (**a**) and 5 (**b**).

Sequential condition model validation

The Sequential condition models also make predictions about the relationship between the diffusion constants of remembered Computed and Perceived values. Once again, we assessed how well participant behavior matched these assumptions, detailed in ***Equation 11*** for AtD and ***Equation 12*** for DtA (***Figure 8***). We fit a line to the measured variances in reporting error as a function of delay for a given set size in both Perceived and Computed Sequential blocks to estimate the change in variance over time (the empirical diffusion constant estimates: $\hat{\sigma}_1^2$, $\hat{\sigma}_{NE}^2$, $\hat{\sigma}_{NL}^2$, $\hat{\sigma}_{MN\text{-}seq}^2$, where N=2 or 5 for the two set sizes). We then compared the difference of these empirical estimates to the predictions of the best-fit model for each participant (***Figure 8***).

In general, the participant data conformed to the model predictions of the best-fit model for each participant, despite substantial individual variability. For participants whose data were best fit by the AtD model (n=9 for both set sizes), the difference between empirical estimates of the diffusion

their model predictions (***Figure 9a, e, c and g***). Conversely, participants' best fit by DtA exhibited diffusion constants that were notably smaller for Computed mean locations versus single Perceived locations (***Figure 9j and l***; lower slope of cyan/blue line vs. purple line). The corresponding average predictions by the best fit DtA models for error variance of Early and Late items also aligned with participant data from DtA fit participants (***Figure 9b, f, d and h***). We also compared the variance in AtD and DtA participants' reports of the mean across delays using an ANOVA and multiple comparisons but found no significant differences in variability at any delay between models ($p>0.05$).

Sequential condition strategy comparisons

Across the population, participants had roughly equal tendencies to use either one of the two strategies (AtD or DtA) for the two set-size conditions (***Figure 0.431 0.431***

multiple quantities stored at once. Third, our DtA model also assumed that each item was stored individually. Alternatively, items could have been discarded or merged (chunked) (***Krishnan et al., 2018***; ***Wei et al., 2012***), leading to different memory loads which could also affect performance. Fourth, most of our participants used strategies that were well described by the AtD or DtA model. However, under certain conditions (i.e., Sequential, set size 5) some participants seemed to use hybrid strategies. This kind of strategy would suggest extensive flexibility in when and how evidence is incorporated into computed decision variables, thereby placing potentially complex demands on working memory.

Both of our primary models were based on assumptions of a drifting memory representation. This random drift is traditionally associated with attractor models of working memory (***Bays, 2014***; ***Compte et al., 2000***; ***Macoveanu et al., 2007***; ***Wei et al., 2012***) that have been used extensively to describe the underlying neural mechanisms (***Funahashi et al., 1989***; ***Shafi et al., 2007***; ***Takeda and Funahashi, 2002***; ***Wimmer et al., 2014***). In these models, neural network activity is induced by an external stimulus and then maintained via excitatory connections of similarly tuned neurons and long-ranged inhibition. Random noise causes the center of this activity (which represents the stimulus) to drift in a manner that, dependent on the implementation, can depend on the delay duration, set size, and/or their interaction (***Almeida et al., 2015***; ***Bays, 2014***; ***Koyluoglu et al., 2017***). A recent implementation even can naturally compute a running average based on sequentially presented information (***Esnaola-Acebes et al., 2021***). Our results imply that such models should be extended to support the flexible use of different strategies that govern when and how incoming information is used to form such averages. It will be interesting to see if such a flexible model can account for neural activity in the dorsolateral prefrontal cortex, which includes neurons with persistent activity that has been associated with both spatial working memory (***Compte et al., 2000***; ***Constantinidis et al., 2018***; ***Riley and Constantinidis, 2016***; ***Wei et al., 2012***; ***Wimmer et al., 2014***) and the formation of decisions based on an accumulation of evidence (***Curtis and D'Esposito, 2003***; ***Heekeren et al., 2006***; ***Heekeren et al., 2008***; ***Kim and Shadlen, 1999***; ***Lin et al., 2020***; ***Philiastides et al., 2011***).

In conclusion, we found that in this spatial, continuous task, participant accuracy for both perceived and computed values was subject to working-memory limitations of both time and capacity. Additionally, we found behavior that was consistent with both the storage strategies we investigated. The fact that different participants employed different strategies for storing a computed value (such as a decision variable) and that these strategies have different consequences on overall accuracy has important implications for not only future neural network models of working memory, but also for future computational models of decision-making.

Materials and methods

Human psychophysics behavioral task

We tested 17 participants (4 males, 12 females, 1 chose not to answer; age range=22–87 years). The task was created with PsychoPy3 (***Peirce et al., 2019***) and distributed to participants via Pavlovia.org, which allowed participants to perform the task on their home computers after providing informed consent. These protocols were reviewed by the University of Pennsylvania Institutional Review Board (IRB) and determined to meet eligibility criteria for IRB review exemption authorized by 45 CFR 46.104, category 2.

Participants were instructed to sit one arm-length away from their computer screens during the experiment and to use the mouse to indicate choices. Each participant completed 1–2 sets of four blocks of trials in their own time.

The basic trial structure is illustrated in ***Figure 1***. Each trial began with the presentation of a central white fixation cross (1% of the screen height). The participant was instructed to maintain fixation on this cross when not actively responding. The participant began each trial by placing the mouse over the cross and clicking, to allow for self-pacing and pseudo-fixation. Initiating a trial caused a white annulus of radius 25% of the screen height to appear. A block-specific memory array appeared 250 ms later, centered at an angle chosen uniformly and at random on the annulus. The array consisted of 1, 2, or 5 colored disks sized 1.5% screen in diameter. The angular difference between any two adjacent disks was at least 6°, and between the two most distal disks was at most 60°. The disks from clockwise to counter-clockwise were always presented in the same order: green, red, blue, magenta,

the effects of set size, delay duration, and task context on response variability using a two-way repeated measures ANOVA. On Simultaneous Perceived and Computed blocks, we used a 3 (delay duration: 0, 1, or 6 s) × 3 (set size: 1, 2, or 5 disks) within-participant design. On Sequential Perceived blocks, we used a 2 (delay duration: 1 or 6 s) × 3 (set size: 1, 2, or 5 disks) within-participants design for stimuli presented at the beginning of the trial (Early) and a 2 (delay: 0.5 or 3 s) × 2 (set size: 2 or 5 disks) design for stimuli presented halfway through the trial (Late). On Sequential Computed blocks, we used a 2 (delay duration: 1 or 6 s) × 3 (set size: 1, 2, or 5 disks) within-participants design. When the comparison included set size=1, data were always taken from the Simultaneous Perceived block.

To assess performance differences based on strategy use, additional analyses were performed once the data had been fit to the models and the best fit model had been selected (see below). These analyses included an assessment of response error variability in the Computed blocks using a 2 (model: AtD or DtA) × 3 or 2 (delay: 0, 2, or 6 s Simultaneous condition, 1 or 6 s for Sequential) ANOVA with multiple comparisons to identify differences. To interrogate best fit parameter differences, two-sided *t*-tests were used to see if the mean difference in best-fit parameter between AtD and DtA participants was significantly different from 0 for both Simultaneous and Sequential conditions. To assess learning effects, a two-sided, paired *t*-test was used to see if the mean or standard deviation of error responses in set size 5 Sequential conditions differed between the first and second half of trials (we found no difference at either delay: for 1 s delay p=0.67 and 0.11 for mean and standard deviation, respectively; for 6 s delay p=0.75 and 0.98 for mean and standard deviation, respectively).

Model-based analyses

Our models were based on principles of working memory that are well described by bump-attractor network models (***Compte et al., 2000***; ***Laing and Chow, 2001***; ***Wimmer et al., 2014***). In such models, stimulus location is represented by a 'bump' in activity from neurons tuned to that and similar locations. These neurons recurrently activate each other, maintaining a bump of activity even after stimulus cessation. However, because of the stochastic nature of neural activity and synaptic transmission (***Faisal et al., 2008***), there is variability in which neurons have the most activity at any given time (and thus are the center of the bump representing the stimulus). This variability in bump center corresponds to variability in the location representation and a degradation of the memory representation over time. The dynamics of this bump can be described as a diffusion process that obeys Brownian motion (***Compte et al., 2000***; ***Kilpatrick, 2018***; ***Kilpatrick et al., 2013***; ***Laing and Chow, 2001***). We used this simplified description in our models as follows.

Perceived values in working memory

A single point (i.e., the central spatial location of a single disk), x_1, is assumed to be represented in working memory by $x_{t,1}$, where t represents the time since the removal of the stimulus. We assume that $x_{t,1}$ evolves like a sample from a Brownian-motion process. Specifically, when x_1 is observed, it is encoded with some perceptual noise, η^p. Therefore, at time zero, $x_{0,1} \sim N(x, \eta^p)$. This representation accumulates noise over time with some diffusion constant, σ_1^2, further degrading the representation of $x_{t,1}$ from x_1 such that $x_{t,1} \sim N(x_1, \eta^p + t*\sigma_1^2)$. There is additional motor noise in the participant's report, $r_{t,1}$, and we denote the variance of this motor noise by η^m. Mathematically, it is equivalent to add the motor noise at the beginning or the end of the diffusion of $x_{t,1}$ when considering the report, $r_{t,1}$. In our model, we thus represent the sum of the perceptual and motor noise as a single, static noise term. Hence, we show simulated trajectories of as(i80n200D>>> BDC 18.659 -00 9 267.2879 2i /MC8 ( )TjEMC ET/P

e

corresponding to the increased memory load. The representation of the Late item then diffuses for only half of the delay time, T (see ***Figure 2d and e***). We formalized this process with the following model for the report error of the Early ($e_{T,NE}$) and Late ($e_{T,NL}$) items:

$$e_{T,NE} \sim N(0, \eta_{NE} + T/2^{*}\sigma_1^2{}^{*}(N-1)^{A} + T/2^{*}\sigma_1^2{}^{*}N^{A})$$

following conditions. Perceived: delays 1, 3, and 6 s; array size 1 (***Equation 3a***). Perceived: delays 3 and 6 s, array size *N* for both Early (***Equation 6a***) and Late (***Equation 6b***) items. Computed: delays 3 and 6 s, array size *N* (***Equation 7*** for AtD or ***Equation 8*** for DtA).

Because the mean error for each individual participant was not always 0, when fitting the AtD and DtA models we used the empirical mean error from the condition being fitted as a fixed bias term in the model. Mean error and CIs for each participant for each condition are shown in ***Figure 3—figure supplements 2 and 3***; ***Figure 3—figure supplements 2 and 3***.

We obtained separate maximum-likelihood fits for AtD and DtA models for each individual participant, using the function fmincon in MATLAB to minimize the summed negative log-likelihood of obtaining the observed errors for a given condition according to the above equations. Initial parameter values were randomized and the fitting repeated to avoid local minima. Because all models within a given condition had the same number of parameters, we compared log-likelihoods to determine the best-fitting model for a given participant. Because the number of parameters is the same, comparing likelihoods produces equivalent model selection to BIC or AIC.

Assessing model assumption and identifiability

To assess how well each participant's data matched the assumptions of the AtD and DtA models, we also fit a line to the variances of response errors across delays for a given condition for a given participant to obtain empirical estimates of the various diffusion constants (e.g., slope of lines in ***Figure 2b***; empirical estimate of a Perceived value,

relationships. Participants whose empirical diffusion constant relationships fell within the central 95% of the simulated expected range were considered well fit by their model.

To assess model identifiability, for each participant and condition, we fit both models to the results of each set of 1000 simulations generated using the best-fitting parameters from the best-fitting model for that participant and condition. We used the log-likelihoods to determine the best model for each simulation and determined the percentage of correctly identified models. We used these

modelsaccoeach pd-

Additional files

S .e e a .e

- Transparent reporting form

Da a a .ab .

All analysis code is available on GitHub (https://github.com/TheGoldLab/Memory_Diffusion_Task, copy archived at swh:1:rev:69cee7449f92f9d19148332979087bf4e6a9f867). Data used for figures will be made available on Dryad.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Schapiro K, Josic K, Gold J, Kilpatrick Z	2022	Memory array locations, delay times, and participant response	https://doi.org/10.5061/dryad.w3r2280rm	Dryad Digital Repository, 10.5061/dryad.w3r2280rm

References

Almeida R, Barbosa J, Compte A. 2015. Neural circuit basis of visuo-spatial working memory precision: a computational and behavioral study. *Journal of Neurophysiology* **114**:1806–1818. DOI: https://doi.org/10.1152/jn.00362.2015, PMID: 26180122

Bastos AM, Loonis R, Kornblith S, Lundqvist M, Miller EK. 2018. Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *PNAS* **115**:1117–1122. DOI: https://doi.org/10.1073/pnas.1710323115, PMID: 29339471

Bays PM, Husain M. 2008. Dynamic shifts of limited working memory resources in human vision. *Science (New York, N.Y.)* **321**:851–854. DOI: https://doi.org/10.1126/science.1158023, PMID: 18687968

Bays PM, Catalao RFG, Husain M. 2009. The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision* **9**:7. DOI: https://doi.org/10.1167/9.10.7, PMID: 19810788

Bays PM. 2014. Noise in neural populations accounts for errors in working memory. *The Journal of Neuroscience* **34**:3632–3645. DOI: https://doi.org/10.1523/JNEUROSCI.3204-13.2014, PMID: 24599462

Bernacchia A, Seo H, Lee D, Wang XJ. 2011. A reservoir of time constants for memory traces in cortical neurons. *Nature Neuroscience* **14**:366–372. DOI: https://doi.org/10.1038/nn.2752, PMID: 21317906

Brady TF, Alvarez GA. 2015. Contextual effects in visual working memory reveal hierarchically structured memory representations. *Journal of Vision* **15**:6. DOI: https://doi.org/10.1167/15.15.6, PMID: 26575192

Brody CD, Hanks TD. 2016. Neural underpinnings of the evidence accumulator. *Current Opinion in Neurobiology* **37**:149–157. DOI: https://doi.org/10.1016/j.conb.2016.01.003, PMID: 26878969

Compte A, Brunel N, Goldman-Rakic PS, Wang XJ. 2000. Synaptic Mechanisms and Network Dynamics Underlying Spatial Working Memory in a Cortical Network Model. *Cerebral Cortex (New York, N.Y* **10**:910–923. DOI: https://doi.org/10.1093/cercor/10.9.910, PMID: 10982751

Constantinidis C, Funahashi S, Lee D, Murray JD, Qi XL, Wang M, Arnsten AFT. 2018. Persistent Spiking Activity Underlies Working Memory. *The Journal of Neuroscience* **38**:7020–7028. DOI: https://doi.org/10.1523/JNEUROSCI.2486-17.2018, PMID: 30089641

Cowan N, Morey CC, Chen Z, Gilchrist AL, Saults JS. 2008. Theory and Measurement of Working Memory Capacity Limits. *Advances in Research and Theory* **9**:49–104. DOI: https://doi.org/10.1016/S0079-7421(08)00002-9

Curtis CE, D'Esposito M. 2003. Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences* **7**:415–423. DOI: https://doi.org/10.1016/s1364-6613(03)00197-9, PMID: 12963473

Esnaola-Acebes JM, Roxin A, Wimmer K. 2021. Bump attractor dynamics underlying stimulus integration in perceptual estimation tasks. *Neuroscience* **10**:434192. DOI: https://doi.org/10.1101/2021.03.15.434192

Faisal AA, Selen LPJ, Wolpert DM. 2008. Noise in the nervous system. *Nature Reviews Neuroscience* **9**:292–303. DOI: https://doi.org/10.1038/nrn2258

Funahashi S, Bruce CJ, Goldman-Rakic PS. 1989. Mnemonic Coding of Visual Space in the Monkey's Dorsolateral Prefrontal Cortex. *Journal of Neurophysiology* **61**:331–349. DOI: https://doi.org/10.1152/jn.1989.61.2.331

Gold JI, Shadlen MN. 2007. The neural basis of decision making. *Annual Review of Neuroscience* **30**:535–574. DOI: https://doi.org/10.1146/annurev.neuro.29.051605.113038, PMID: 17600525

Gold JI, Stocker AA. 2017. Visual Decision-Making in an Uncertain and Dynamic World. *Annual Review of Vision Science* **3**:227–250. DOI: https://doi.org/10.1146/annurev-vision-111815-114511, PMID: 28715956

Heekeren HR

Kilpatrick ZP, Ermentrout B, Doiron B. 2013. Optimizing working memory with heterogeneity of recurrent cortical excitation. *Journal of Neuroscience* **33**:18999–19011. DOI: https://doi.org/10.1523/JNEUROSCI.1641-13.2013, PMID: 24285904

Kilpatrick ZP. 2018. Synaptic mechanisms of interference in working memory. *Scientific Reports* **8**:1–20. DOI: https://doi.org/10.1038/s41598-018-25958-9, PMID: 29777113

Kim JN, Shadlen MN. 1999. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience* **2**:176–185. DOI: https://doi.org/10.1038/5739, PMID: 10195203

Koyluoglu OO, Pertzov Y, Manohar S, Husain M, Fiete IR. 2017. Fundamental bound on the persistence and capacity of short-term memory stored as graded persistent activity. *eLife* **6**:e22225. DOI: https://doi.org/10.7554/eLife.22225

Krishnan N, Poll DB, Kilpatrick ZP. 2018. Synaptic efficacy shapes resource limitations in working memory. *Journal of Computational Neuroscience* **44**:273–295. DOI: https://doi.org/10.1007/s10827-018-0679-7, PMID: 29546529

Laing CR, Chow CC. 2001. Stationary bumps in networks of spiking neurons. *Neural Computation* **13**:1473–1494. DOI: https://doi.org/10.1162/089976601750264974, PMID: 11440594

Lin Z, Nie C, Zhang Y, Chen Y, Yang T. 2020. Evidence accumulation for value computation in the prefrontal cortex during decision making. *PNAS* **117**:30728–30737. DOI: https://doi.org/10.1073/pnas.2019077117, PMID: 33199637

Liu ASK, Tsunada J, Gold JI, Cohen YE. 2015. Temporal integration of auditory information is invariant to temporal grouping cues. *ENeuro* **2**:ENEURO.0077-14.2015. DOI: https://doi.org/10.1523/ENEURO.0077-14.2015, PMID: 26464975

Macoveanu J, Klingberg T, Tegnér J. 2007. Neuronal firing rates account for distractor effects on mnemonic accuracy in a visuo-spatial working memory task. *Biological Cybernetics* **96**:407–419. DOI: https://doi.org/10.1007/s00422-006-0139-8, PMID: 17260154

Oberauer K, Farrell S, Jarrold C, Lewandowsky S. 2016. What limits working memory capacity *Psychological Bulletin* **142**:758–799. DOI: https://doi.org/10.1037/bul0000046, PMID: 26950009

Panichello MF, DePasquale B, Pillow JW, Buschman TJ. 2019. Error-correcting dynamics in visual working memory. *Nature Communications* **10**:3366. DOI: https://doi.org/10.1038/s41467-019-11298-3, PMID: 31358740

Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Kastman E, Lindeløv JK. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* **51**:195–203. DOI: https://doi.org/10.3758/s13428-018-01193-y, PMID: 30734206

Philiastides MG, Auksztulewicz R, Heekeren HR, Blankenburg F. 2011. Causal role of dorsolateral prefrontal cortex in human perceptual decision making. *Current Biology* **21**:980–983. DOI: https://doi.org/10.1016/j.cub.2011.04.034, PMID: 21620706

Ploner CJ, Gaymard B, Rivaud S, Agid Y, Pierrot-Deseilligny C. 1998. Temporal limits of spatial working memory in humans. *European Journal of Neuroscience* **10**:794–797. DOI: https://doi.org/10.1046/j.1460-9568.1998.00101.x, PMID: 9749746

Ratcliff R, Smith PL, Brown SD, McKoon G. 2016. Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences* **20**:260–281. DOI: https://doi.org/10.1016/j.tics.2016.01.007, PMID: 26952739

Riley MR, Constantinidis C. 2016. Role of Prefrontal Persistent Activity in Working Memory. *Frontiers in Systems Neuroscience* **9**:181. DOI: https://doi.org/10.3389/fnsys.2015.00181, PMID: 26778980

Schneegans S, Bays PM. 2018. Drift in Neural Population Activity Causes Working Memory to Deteriorate Over Time. *The Journal of Neuroscience* **38**:4859–4869. DOI: https://doi.org/10.1523/JNEUROSCI.3440-17.2018, PMID: 29703786

Shadlen MNN, Shohamy D. 2016. Decision Making and Sequential Sampling from Memory. *Neuron* **90**:927–939. DOI: https://doi.org/10.1016/j.neuron.2016.04.036, PMID: 27253447

Shafi M, Zhou Y, Quintana J, Chow C, Fuster J, Bodner M. 2007. Variability in neuronal activity in primate cortex during working memory tasks. *Neuroscience* **146**:1082–1108. DOI: https://doi.org/10.1016/j.neuroscience.2006.12.072, PMID: 17418956

Document (/10.1523/JNEUROSCI.3440-17.2018)Tj 0 -1 Td (Schneegans S, Bays PM. 2018. Drift in Neural Population Activity Causes Working Memory to Deteriorate Over Time. The Journal of Neuroscience 38:4859–4869. DOI: https://doi.org/10.1523/JNEUROSCI.3440-17.2018, PMID: 29703786)Tj 0.431 0.431 0.439

Wimmer K, Nykamp DQ, Constantinidis C, Compte A. 2014. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nature Neuroscience*