

## **UC Merced**

# **Proceedings of the Annual Meeting of the Cognitive Science Society**

### **Title**

People Adjust Recency Adaptively to Environment Structure

### **Permalink**

<https://escholarship.org/uc/item/9tt7193t>

### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

### **ISSN**

1069-7977

### **Authors**

Luthra, Mahi K  
Todd, Peter M.

### **Publication Date**

2021

Peer reviewed

# People Adjust Recency Adaptively to Environment Structure

**Mahi Luthra (mkluthra@iu.edu)**

Cognitive Science Program, Indiana University Bloomington (IUB);  
Department of Psychological and Brain Sciences, IUB  
1001 E 10th St. Bloomington, IN 47405 USA

**Peter M. Todd (pmtodd@indiana.edu)**

Cognitive Science Program, IUB;  
Department of Psychological and Brain Sciences, IUB  
1101 E 10th St. Bloomington, IN 47405 USA

## Abstract

Recency effects—giving exaggerated importance to recent outcomes—are a common aspect of decision tasks. In the current study, we explore two explanations of recency-based decision making, that it is (1) a deliberate strategy for adaptive decision making in real-world environments which tend to be dynamic and autocorrelated, and/or (2) a product of processing limitations of working memory. Supporting explanation 1, we found that participants strategically adjusted their recency levels across trials to achieve optimal levels in a range of tasks. Furthermore, they started with default recency values that had high aggregate performance across environments. However, only some correlations between recency values and WM scores were significant, providing no clear conclusion regarding explanation 2. Ultimately, we propose that recency involves a combination of the two—people can strategically change recency within the limits of WM capacities to adapt to external environments.

**Keywords:** recency; decision making; working memory

## Introduction

Humans (and other animals) often give exaggerated importance to small samples of recent information. The existence of such recency effects was initially demonstrated through short-term memory recall tasks and was attributed to retrieval of the last few list items from the item-limited short-term memory store (Murdock, 1962). However, this initial short-term memory conceptualization and explanation was challenged as researchers found recency effects even in long-term memory tasks, where information had to be retained across long stretches of time intervals interleaved with distractor tasks (Bjork & Whitten, 1974).

Since these initial studies with memory recall, a wider range of cognitive paradigms, such as decision-making tasks, have demonstrated recency effects. For instance, in sequential decision making (e.g., choosing among multiple gambles repeatedly), researchers find that participants choose alternatives that are successful in recent trials (da Silva et al., 2017). The ubiquity of recency effects in decision making is particularly evident in the wide-spread inclusion of a recency weighting parameter that discounts the influence of older outcomes in many reinforcement learning models (e.g., EVL and PVL models; Busemeyer & Stout, 2002; Erev & Roth, 1998). In fact, researchers find that extremely simplistic last-

outcome-based win-stay-lose-shift models also fit human decision data successfully, highlighting the prevalence of recency in decision making (Worthy & Maddox, 2014).

In the current study, we are interested in exploring explanations for recency-based decision-making. We compare two potential explanations, that it is either (1) a deliberate strategy for adaptive decision making in real-world environments, or (2) a product of computational limitations.

## Recency as a Deliberate Decision Strategy

Most natural environments are dynamic and autocorrelated, both across time and space (e.g., fertility of land, cultural customs). Recency is likely to be an advantage in such environments where a small set of recent experiences is usually representative of the current state of the world. Here, averaging instead across a large pool of previous experiences can obscure useful information. Recency as an adaptation to real world distributions can be observed across animal species. For instance, Real (1992) studied foraging behavior of bees, finding that they made search decisions based on a relatively small number of visits (approximately three)—an effective strategy for the highly spatially and temporally autocorrelated nectar distribution in their natural habitat. Similarly, Anderson and Schooler (1991) found evidence for temporal autocorrelation in informational environments that humans regularly interact with (word frequencies in the New York Times, emails, and parental speech). They argued that stronger recall for recent events serves as an adaptation to these external patterns, making the things that are likely to reoccur more readily recallable from memory.

Recency has thus been proposed as an adaptive strategy for decision making. Decision strategies may contain learning mechanisms that enable flexible updating of parameters so that they become more successful in the current environmental structure (Gallistel, 1990). For instance, environmental challenges such as deciding when to give up seeking nectar from flowers can vary in how much to rely on recency over time (for instance, from flower patch to flower patch). For organisms facing such environments reliably, a flexible strategy that can adjust its use of recency could be more advantageous, and hence more likely to evolve, than an inflexible recency strategy. Supporting this intuition, Wilke and Barrett (2009) found that recent-outcome-based decision

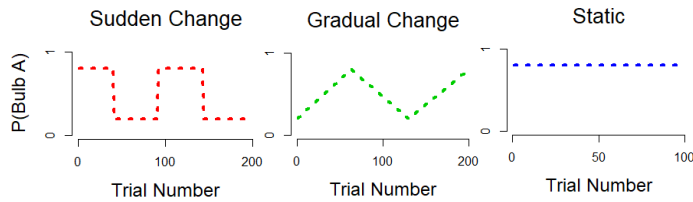


Figure 1: Decision task structures.

making was a default strategy for their (human) participants, but it could be reduced with experience in random environments (e.g., predicting coin tosses) where it was not useful. In line with that work, we examine the default recency used by participants and their ability to flexibly tune it to suit the decision structure provided.

### Recency Produced by Cognitive Limitations

Another possible explanation for the prevalence of small sample recency in an agent’s behavior is that it arises from inescapable limitations in the agent’s cognitive abilities. Humans can maintain and manipulate limited amounts of information at one time—this computationally limited capacity is commonly referred to as working memory (WM; Cowan, 2017). As stated earlier, researchers initially conceived of recency as a product of retrieval from the short-term store of the WM. Although this conceptualization does not sufficiently explain recency effects found in tasks involving long-term memory, it is possible that WM limitations are also responsible for long-term recency since they make it expensive to sample and process large amounts of information for decision making simultaneously. Some researchers suggest that the experience of mental effort or fatigue with extensive information processing is a result of sampling information from long-term memory and processing it in WM, producing opportunity costs for the information included in deliberations (Shenhav et al., 2017; Vul et al., 2014). Since recent information is often more valuable than older, humans needing to limit to small samples of information might retrieve recent information preferentially, and thus make decisions giving exaggerated importance to it.

Such an account would suggest that recency effects are contingent on WM limitations, and that smaller WM capacity could lead to stronger reliance on more recent information. Several studies have attempted to investigate relationships between individual differences in recency and working memory, obtaining inconclusive results. For instance, Luthra and Todd (2019) found weak correlations ( $p < .1$ ) of recency with WM storage measures. Ashby and Rakow (2014) obtained stronger correlations, but further analysis by Wulff and Pachur (2016) found that their models had low discriminability and potentially less reliable estimates of recency. In the current study, we compare modelled recency parameters across task structures to provide more conclusive results on the relationship between recency and WM.

### Current Study

We have provided two conceptualizations of recency—as a strategy adapted to typical decision environments and as a bias resulting from cognitive limitations. These explanations are not mutually exclusive. For instance, in some environments (where the optimal sample size for recency-based decision making is smaller than WM size), participants might strategically choose to engage in recency, and in others (where a larger sample is optimal or where humans are already mentally overloaded) recency might be produced by WM constraints.

The current study explores these alternative explanations for recency (strategy, bias, or both) in a probability learning task. This task entails sequential trials of choosing between two mutually exclusive and exhaustive outcomes. For example, in one setting participants are presented with two lightbulbs—Bulb A and Bulb B with 0.8 and 0.2 probabilities of turning on respectively—and on each trial, participants must predict which lightbulb will turn on without knowing these underlying probabilities. Researchers find that responses in this type of probability learning task are frequently influenced by recent outcomes (da Silva et al., 2017). Under a static task structure (where probabilities of lightbulbs remain constant across trials), recency would be a comparatively poor strategy since it would produce noisy calculations of expected value from small samples of only recent events. On the other hand, under a changing and autocorrelated task structure, recency would be a more effective strategy, since only a small recent sample would be representative of the current expected value.

In our current experiment, each participant was given one of two task structures—gradual change (lightbulb probabilities changed gradually across time) and sudden change (there were sudden shifts in probabilities). We compared our data with results from Luthra and Todd (2019) where probabilities of lightbulbs remained constant across time. This gives three task structures to consider—sudden change, gradual change, and static (displayed in Figure 1). These were chosen for two reasons. First, the two changing task structures represent aspects of particular real-world environments (which will often be more complex) where events are autocorrelated and they either change gradually (e.g., nectar content of flowers) or suddenly (e.g., appearance of a predator). Second, the level of recency that is most effective in each of the three environments differs—extreme recency is adaptive in sudden change, less recency is adaptive in gradual change, and recency is entirely maladaptive in the static environment. The varying usefulness of recency in the three environments is shown in Figure 2 which plots results of simulations comparing performance of different values of recency (defined as amount of temporal discounting—see Modelling section below).

In our experiments we aim to determine whether participants can tune their recency to appropriate levels for each structure, providing support to the conceptualization of recency as strategic. We also correlate recency values with WM—high correlations between the two, and similar recency

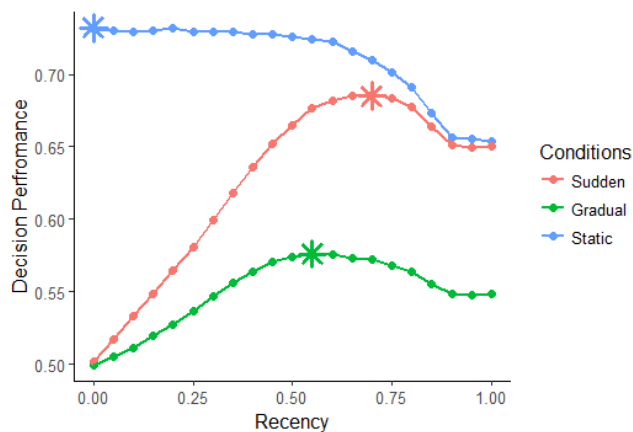


Figure 2: Performance of recency parameter values across task structures. Asterisk indicates optimal recency parameter.

values across tasks (rather than adaptive tuning to each task), would support the conceptualization of recency as a computational bias. As mentioned earlier, participant behavior might be produced by both combined—here, we could find that in task structures calling for high recency (i.e., sudden change) participants strategically reduce their sample size, and in task structures calling for no recency (i.e., static environments) sample size is correlated with WM limits.

## Methods

The current experiment with gradual and sudden change task structures and the earlier experiment with the static structure (Luthra & Todd, 2019) used very similar methods, so the following description applies to both. We have specified any divergences in procedures.

## Participants

Sixty-five undergraduate students of Indiana University served as participants for this study. Of these, data of four participants had to be discarded due to failure to complete at least one of the tasks. This was a between-groups study—34 and 31 participants were given the sudden and gradual change conditions respectively. For the static task structure, data of 48 participants (out of 123 total) was used from Luthra and Todd (2019), as explained below.

## Procedures and Tasks

In both studies, participants were provided with five computer-based tasks in the laboratory, beginning with four tasks measuring WM capacity, followed by the probability learning task.

**Memory tasks.** Participants performed the WM tasks in the following order: symmetry span, digit span, visual array, and operation span. The digit span is a classic short-term memory task involving simple number recall (method similar to Woods et al., 2011). The visual array task involves identifying rapid color changes in arrays of 4, 6, 8, or 10

colored squares (method similar to Cowan et al., 2006). Symmetry and operation span involve short-term retention of information (spatial positions of colored squares for symmetry span and letters for operation span) across intervals of distractor processing tasks (symmetry or arithmetic accuracy judgements respectively; Oswald et al., 2015).

These specific WM tasks were selected because they involve usage of functionally different WM components—digit span and visual array use WM storage; symmetry and operation span use WM control (see Cowan, 2008 for difference between the two). They also use different content modalities—digit and operation span are verbal numeric; and symmetry span and visual array are visuospatial.

**Probability learning task.** In the probability learning task participants were presented with a ‘bulb-box’ on the computer screen—a device containing two lightbulbs (Bulb A and Bulb B). They were informed that one of the two bulbs would turn on in each trial (across a total of 200 trials) and they had to choose the correct bulb. For correct guesses, participants gained a point and for incorrect guesses, they lost a point. To motivate participants to perform optimally, they were awarded with 0 to 3 nutritional bars depending on performance.

In the sudden change condition, a randomly chosen bulb would start initially with 0.8 probability of being correct. After several trials, this probability would suddenly drop to 0.2 (the probability of the other bulb increased to 0.8). Across 200 trials, this flipping of optimal lightbulb occurred three times (between trials 40-60, 90-110, and 140-160). In the gradual shift condition, the bulbs changed probabilities gradually across time, with the optimal bulb flipping gradually at similar timepoints. The static condition data from Luthra and Todd (2019) had only 100 trials, but each participant performed the task thrice, each with a different probability contingency—0.6, 0.7, and 0.8—in a random order. For the current analyses, we only used data from the 48 participants who saw the 0.8 probability contingency (to match the probability of the present experiment) in the first task (preventing any practice effects). Figure 1 displays examples of all three conditions.

## Models

The collected data was fit with a recency and a random Bernoulli model. We assumed that participants who were better fit by random Bernoulli were using an alternate unaccounted strategy (e.g., random guessing, pattern matching) while participants who were better fit by the recency model were using some variation of recency. Therefore, we estimated the frequency of recency in the sample (in comparison to random Bernoulli) and by fitting the recency parameter, we estimated the strength of recency.

The recency model is a variant of the expectancy-valence learning (EVL) model frequently used to fit decision making data (Busemeyer & Stout, 2002). It has two parameters—recency ( $R$ ) and exploration ( $c$ ). The model assumes that

participants assign a utility value  $u$  on every trial  $t$  to each bulb  $j$  based on the current outcome  $x$  of the ‘bulb-box’:

$$u_j(t) = \begin{cases} 1 & \text{if } x(t) = j, \\ 0 & \text{if } x(t) \neq j \end{cases}$$

Therefore, if the current outcome is  $j$ , bulb  $j$  is given utility of 1, and otherwise 0. Across trials, this utility is incorporated into the running expected utility  $E_j$  of both bulbs while discounting the value of older outcomes in proportion to the recency parameter  $R$ :

$$E_j(t) = \left(1 - \left(\frac{1}{t}\right)^{1-R}\right) \cdot E_j(t-1) + \left(\frac{1}{t}\right)^{1-R} \cdot u_j(t)$$

Here,  $R$  can vary from 0 (no recency, averaging of all outcomes) to 1 (extreme recency, only last outcome included in calculations; Hertwig et al., 2006). On each trial, the expected utility  $E_j$  produces decisions  $D$  through a softmax version of Luce’s choice rule which incorporates exploration:

$$\Pr[D(t+1) = j] = \frac{e^{\theta(t) \cdot E_j(t)}}{\sum_{k=1}^2 e^{\theta(t) \cdot E_k(t)}}; \quad \theta(t) = \left(\frac{t}{20}\right)^c$$

Here, exploration  $c$  can range from 0 (high exploration) to 1 (high exploitation based on calculated utility). In the above equation, the  $\left(\frac{t}{20}\right)^c$  term leads the model to gradually move from exploration to exploitation across trials and fits data better than a model assuming constant exploration. We incorporated exploration in our model even though it is irrelevant to our hypothesis because the model fit the data better with its addition (despite punishing for increased degrees of freedom), indicating that exploration was important to participant responding.

The second model was a simple random Bernoulli model with only one parameter—the probability that participants choose bulb  $j$ . Therefore, this model predicts unequal probabilities of choosing the two bulbs, independent of observed outcomes. As mentioned earlier, this model is useful for identifying participants who are using an unaccounted strategy other than recency.

We used Bayesian hierarchical modelling for parameter fitting and model comparison, implemented on JAGS via R. We drew 20,000 samples using three MCMC chains. A categorical distribution was employed to determine the best fitting model for each participant—on each MCMC timestep, one or the other model was sampled in proportion to their probabilities of being the true model. Diagnostic plots and effective sample size indicated appropriate convergence for most parameters. Prior to fitting participant data, we performed parameter recovery on simulated data to ensure that modelling was able to appropriately estimate underlying parameters.

## Results

Overall, the recency model provided a good fit to data and outperformed the random Bernoulli model. The posterior probability  $P(\text{model}=\text{recency}|D)$  was 0.79, Bayes factor was 3.76, and 95 out of 113 participants across the tasks were better fit by the recency model. Since Bayesian model fitting automatically compensates for model complexity (Kruschke,

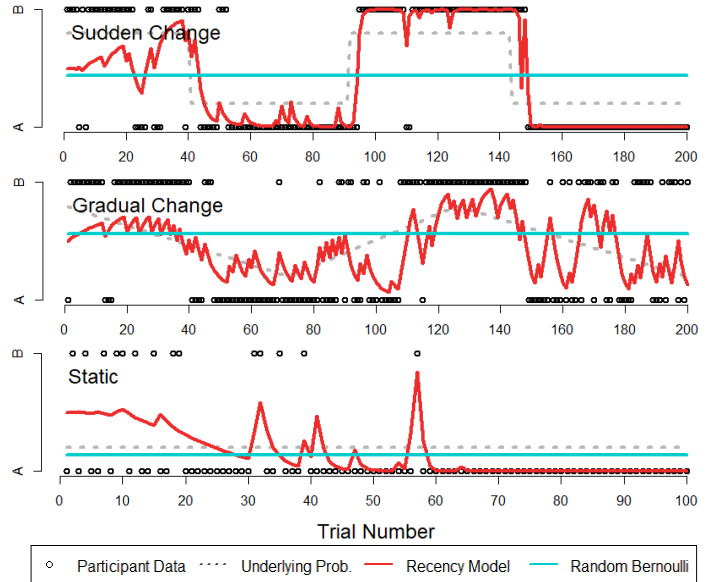


Figure 3: Model fits to individual participants.

2014), we can be confident that the greater performance of the recency model over Bernoulli is not because of its greater degrees of freedom, but because participants made choices fitting a recency approach to decision making. Figure 3 displays model fits to individual participant data.

Estimates of  $R$  obtained from the recency model were then used to analyze recency as an adaptive decision strategy and as a computational bias. Here, we present three analyses: (1) We compared the use of recency by participants across task structures to determine if participants could, through experience, strategically adjust their recency to optimal levels for their task (as displayed in Figure 2). (2) We analyzed change in recency across trials within tasks—this helped us determine the default recency used by participants (in the beginning of the task) and observe how quickly participants adjusted it to suit current task structure. (3) We calculated correlations between recency and WM measures to determine if recency strategy was a result of computational limitations.

### Analysis 1: Recency across Task Structures

Figure 4 displays distribution of recency values across the three task structures. As can be seen, participants used lowest recency for the static task structure ( $M = 0.42$ ;  $SD = 0.12$ ), higher recency for gradual change ( $M = 0.62$ ;  $SD = 0.07$ ) and highest for sudden change ( $M = 0.78$ ;  $SD = 0.04$ ). The mean recency values used roughly correspond with the optimal values displayed in Figure 2. The difference in recency for all three task structure comparisons was significant ( $p < .05$ ), suggesting that participants strategically tuned their recency to the current task structure.

### Analysis 2: Recency across Trials

To quantify the change in recency across trials, we divided participant data into 50-trial phases and fit a different recency parameter to each phase. Therefore, for the sudden and gradual change conditions, we obtained four recency values,

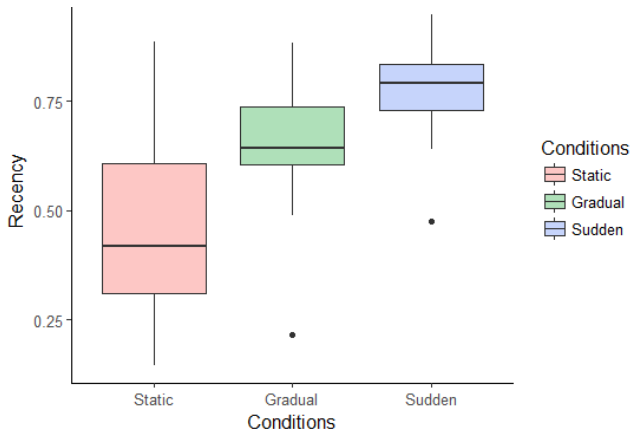


Figure 4: Distribution of recency values across task structures.

each for a different 50-trial phase across the 200 trials. Since data for the static condition from Luthra and Todd (2019) consisted of only 100 trials, these were fit by just two recency values. Prior to modelling, we used parameter recovery on simulated data and confirmed that 50 trials were enough for the model to accurately converge on parameter values.

Figure 5 displays the fitted recency values across trial phases. Across all three task structures, recency in the first 50 trials is similar (no significant difference). The average starting value (possibly a default recency value) is 0.52 ( $SD = 0.17$ ). From this initial starting value, recency parameters in the gradual and sudden change condition are adjusted by participants across trials to values closer to the optimal for those task structures (shown in Figure 2). However, for the static condition, the recency parameter values do not change. There is a significant difference in recency values of the last trial phase between the three conditions ( $p < .05$ ).

### Analysis 3: Correlations between Recency and WM

Table 1 displays correlations between mean recency parameters from Analysis 1 and WM scores. Overall, the

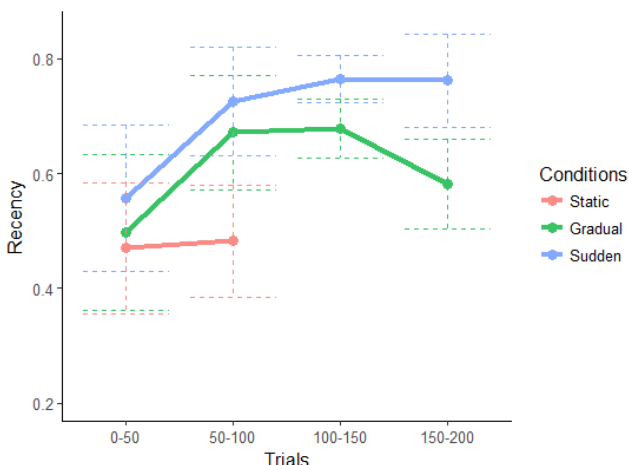


Figure 5: Change in recency across trials for each task structure.

Table 1: Correlations between WM and recency

	Visual Array	Digit Span	Symmetry Span	Operation Span
Sudden	-0.16	-0.32 <sup>+</sup>	-0.08	-0.03
Gradual	-0.1	-0.15	-0.26	-0.13
Static	-0.25 <sup>+</sup>	-0.24 <sup>+</sup>	-0.29 <sup>*</sup>	-0.7

correlations seem weak and mostly not significant. Correlations are highest for the static condition and significant for the symmetry span WM task. Recency values from just the last 50 trials of the tasks provided similar correlation results.

## Discussion

This study explores the use of recency as (1) a decision strategy that can be flexibly adapted to the environment and/or (2) a bias resulting from computational limitations. Results from Analysis 1 suggest that participants successfully tuned their recency strategy to adaptive levels for the task structure they were provided. We investigated this more closely through Analysis 2, where participant data was split into 50-trial phases and a different recency parameter was fit to each phase. Across all task structures, participants started with similar recency values. In gradual and sudden change environments, these were updated over trials to near-optimal levels (seen in Figure 2). This provides support for the conceptualization of recency as a decision strategy that can be flexibly adapted toward optimality in the environment.

However, recency values remained constant for the static condition, although optimal performance here is achieved at 0 recency (i.e., averaging over all outcomes seen). Two possible explanations for this stand out. First, even though optimal performance is achieved at 0 recency, performance is close to plateauing at values below 0.5 (see Figure 2). Potentially, participants compromised slightly on performance in exchange for lesser mental effort (i.e., having to process just a small recent sample rather than averaging over all outcomes). This suggests that if rewards for using 0 recency were higher, participants would engage more mental resources to include larger samples in their decision making. Such an account is supported by Engle (2010), who finds that WM control is state-dependent, being influenced by factors such as motivation, exhaustion, etc. Second, it is possible that there are strict WM limits preventing recency values below 0.5 which process larger samples. Both these explanations would predict correlations between WM and recency scores (possibly, the latter explanation predicts stronger correlations than the former). However, correlation scores were not entirely conclusive—only one WM test had a significant correlation ( $p < .05$ ) with recency values of the static condition while two others had  $p < .1$ . For gradual and sudden change conditions, recency values had non-significant correlations with WM scores, potentially because the optimal sample sizes for decision making were smaller than the typical WM size and were therefore not limited by it. Overall, our results do not negate the possibility of a combined influence of

strategy and computational limits on recency—potentially, participants strategically reduced recency values when optimality was achievable (in sudden and gradual change) and maintained fixed values (limited by WM) when larger samples were needed for optimality (in static task structure). Further studies are needed to investigate the relationship between recency and WM more conclusively.

On average, participants tended to start tasks around recency value 0.52 (Figure 5), which could potentially be a default strategy. Such a default would be useful if it was near optimal across a wide range of environmental structures that an individual is likely to encounter. To investigate this possibility, we simulated static, sudden change, and gradual change environments across a wide range of probability contingencies (0.9, 0.8, 0.7, 0.6) and estimated performance of a range of recency values in these environments. We then scaled performance of each environment to lie between 0 and 1 and aggregated this scaled performance (Figure 6). Here, we see that the aggregate performance across these environments peaks at recency value near 0.6. This is close to the starting recency value we found in our data, which may thus be used because of its adaptivity across environments.

Our results on recency align with previous work in the area. Wilke and Barrett (2009) found that participants had a default positive dependency to previous outcomes—if a coin landed on heads, participants indicated that it was more likely to land on heads again—which could be reduced with experience. According to Wilke and Barrett, this default response is an evolutionary adaptation to foraging in autocorrelated environments—an explanation supported by findings of similar behavior in rhesus monkeys (Blanchard, Wilke, & Hayden, 2014).

Previous work on changing decision environments have used approximations to Bayesian solutions (Brown & Steyvers, 2009; Piray & Daw, 2020) to model human behavior. Since Bayesian analyses are often computationally intractable, these studies assume that humans approximate the posterior Bayesian distribution of decision outcomes by sampling from it by using MCMC methods like particle filtering. These models differ from the non-Bayesian model described in the current work—here, we suggest that participants take small samples from a likelihood distribution

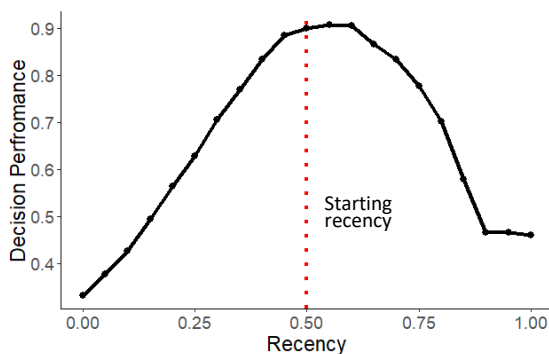


Figure 6: Average performance of recency values across decision environments

consisting of previous experiences (giving preferentiality to recent ones) while Bayesian models suggest that participants store a posterior distribution of hypotheses of the current state of the world (generated from the likelihood distribution and priors) and take small samples from that distribution when making predictions. Incorporating recency in such a particle filter model entails tweaking its learning rate parameter which determines how quickly a posterior distribution changes based on recent experience—such a model was used recently by Piray and Daw (2020). Future analyses can compare the model used here with such Bayesian models.

The current study focuses on a one-way impact of environments on decision making. However, it is more likely that the two mutually interact with each other in a dynamic fashion. That is, human environments influence human decisions, and are in turn influenced by them. For instance, the temporal autocorrelation studied by Anderson and Schooler (1991) in human language-based environments could both be a cause of and a result of recency in decision making—autocorrelation in environments encourages recency-based strategies in decision making which in turn produce autocorrelation in human environments. Dynamical interactions between behavior and environments provide an interesting open direction for decision-making research.

## Data Availability

The simulation code and data are publicly available at <https://github.com/mahiluthra/recency-decisions>

## Acknowledgements

We thank the Adaptive Behavior and Cognition Lab (West) and the Percepts and Concepts Labs for useful feedback on the paper. This research was supported in part by the John Templeton Foundation grant, “What drives human cognitive evolution”.

## References

- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6), 396–408.
- Ashby, N. J. S., & Rakow, T. (2014). Forgetting the past: Individual differences in recency in subjective valuations from experience. *Journal of Experimental Psychology: Learning Memory and Cognition*, 40(4), 1153–1162.
- Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, 6, 173–189.
- Blanchard, T. C., Wilke, A., & Hayden, B. Y. (2014). Hot-hand bias in rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(3), 280–286.
- Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, 58, 49–67.
- Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment:

- Decomposing performance on the Bechara gambling task. *Psychological Assessment*, 14(3), 253–262.
- Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338.
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin and Review*, 24(4), 1158–1170.
- Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Sauls, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition*, 34(8), 1754–1768.
- da Silva, C. F., Victorino, C. G., Caticha, N., & Baldo, M. V. C. (2017). Exploration and recency as the main proximate causes of probability matching: a reinforcement learning analysis. *Scientific Reports*, 7(1), 15326.
- Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Current Anthropology*, 51(1), S17–S26.
- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4), 848–881.
- Gallistel, C. R. (1990). *Learning, development, and conceptual change. The organization of learning*. The MIT Press.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). The role of information sampling in risky choice. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 75–91). New York, NY: Cambridge University Press.
- Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, BUGS, and Stan* (2nd ed.). Cambridge: Academic Press.
- Luthra, M., & Todd, P. M. (2019). Role of working memory on strategy use in the probability learning task. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society* (pp. 721–728). Montreal, QB: Cognitive Science Society.
- Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64(5), 482–488.
- Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, 47(4), 1343–1355.
- Piray, P., & Daw, N. D. (2020). A simple model for learning in volatile environments. *PLOS Computational Biology*, 16(7), e1007963.
- Real, L. A. (1992). Information processing and the evolutionary ecology of cognitive architecture. *American Naturalist*, 140, S108–S145.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, 124, 40–99.
- Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, 38(4), 599–637.
- Wilke, A., & Barrett, H. C. (2009). The hot hand phenomenon as a cognitive adaptation to clumped resources. *Evolution and Human Behavior*, 30(3), 161–169.
- Woods, D. L., Kishiyama, M. M., Lund, E. W., Herron, T. J., Edwards, B., Poliva, O., ... Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology*, 33(1), 101–111.
- Worthy, D. A., & Maddox, W. T. (2014). A comparison model of reinforcement-learning and win-stay-lose-shift decision-making processes: A tribute to W.K. Estes. *Journal of Mathematical Psychology*, 59(1), 41–49.
- Wulff, D. U., & Pachur, T. (2016). Modeling valuations from experience: A comment on Ashby and Rakow (2014). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(1), 158–166.