

Working-Memory Load and Temporal Myopia in Dynamic Decision Making

Darrell A. Worthy
Texas A&M University

A. Ross Otto and W. Todd Maddox
University of Texas at Austin

We examined the role of working memory (WM) in dynamic decision making by having participants perform decision-making tasks under single-task or dual-task conditions. In 2 experiments participants performed dynamic decision-making tasks in which they chose 1 of 2 options on each trial. The decreasing option always gave a larger immediate reward but caused future rewards for both options to decrease. The increasing option always gave a smaller immediate reward but caused future rewards for both options to increase. In each experiment we manipulated the reward structure such that the decreasing option was the optimal choice in 1 condition and the increasing option was the optimal choice in the other condition. Behavioral results indicated that dual-task participants selected the immediately rewarding decreasing option more often, and single-task participants selected the increasing option more often, regardless of which option was optimal. Thus, dual-task participants performed worse on 1 type of task but better on the other type. Modeling results showed that single-task participants' data were most often best fit by a win-stay, lose-shift (WSLS) rule-based model that tracked differences across trials, and dual-task participants' data were most often best fit by a Softmax reinforcement learning model that tracked recency-weighted average rewards for each option. This suggests that manipulating WM load affects the degree to which participants focus on the immediate versus delayed consequences of their actions and whether they employ a rule-based WSLS strategy, but it does not necessarily affect how well people weigh the immediate versus delayed benefits when determining the long-term utility of each option.

Keywords: decision making, working memory load, dual task, reinforcement learning, win-stay lose-shift

Decision making is a recurring task in our everyday lives. Our decisions have both immediate and delayed consequences, and understanding the mechanisms that affect decision making is of much importance. One thing required of the decision maker is adequate cognitive resources to balance the strengths and weaknesses of each choice option. Many common decision-making paradigms used in the laboratory involve participants repeatedly choosing from more than one option and receiving rewards or punishments, often in the form of points or money gained or lost (e.g., Bechara, Damasio, Damasio, & Anderson, 1994; Otto, Markman, Gureckis, & Love, 2010; Rakow & Newell, 2010; Worthy, Maddox, & Markman, 2007; Yechiam, Bussemeyer, Stout, & Bechara, 2005). In these paradigms, the decision maker's goal is

typically to maximize gains and/or minimize losses. Each option must be assigned some expected reward value, and the decision maker must consider not only the immediate benefit that results from selecting each option but also each option's delayed benefits when determining the long-term utility of each option.¹

Some research suggests that development and maintenance of expected reward values for the various choice options require working memory (WM) resources (Bechara & Martin, 2004; Curtis & Lee, 2010; Dretsch & Tipples, 2008; Hinson, Jameson, & Whitney, 2002). Under this view, reduced WM resources might hinder one's ability to develop and maintain expected reward values, which are needed to make the best decisions. One decision-making process that could be dependent on WM resources is simply remembering the rewards given when each option was last chosen. The decision maker can gain valuable insight into the values of each option by simply considering their immediate payoffs.

This article was published Online First April 30, 2012.

Darrell A. Worthy, Department of Psychology, Texas A&M University; A. Ross Otto and W. Todd Maddox, Department of Psychology, University of Texas at Austin.

This research was supported by start-up funds from Texas A&M University to Darrell A. Worthy and by National Institute of Mental Health Grant MH077708 to W. Todd Maddox. A. Ross Otto was supported by a Mike Hogg Endowment Fellowship. We thank Anna Anthony, Corey L. Dossey, Susan Garcia, Karla S. Gomez, Cara Henis, Abigail E. Moser, Crina Silasi-Mansat, Nicole A. Toomey, and Angelina M. Torres for valuable assistance in collecting the data.

Correspondence concerning this article should be addressed to Darrell A. Worthy, Department of Psychology, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235. E-mail: worthyda@tamu.edu

¹ In our paradigm, which involves decision making under uncertainty, the term *expected reward value* is used to represent an estimate of the reward that will be received upon selecting each option. This is distinct from the term *expected value*, which has often been used to represent an optimal objective representation of the benefit of an option. Expected value is calculated by multiplying the reward associated with each option by the probability of receiving the reward, when both are known, as in description-based decision-making paradigms (e.g., Tversky & Kahneman, 1981).

However, another possibility is that people can implicitly learn which options lead to the largest immediate rewards without explicitly remembering the rewards associated with each option. There is much evidence that suggests that a *prediction error*, which is the difference between the reward received and the expected reward value for a given option, is tracked by the ventral striatum, a subcortical region implicated in implicit, procedural learning (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Pagnoni, Zink, Montague, & Berns, 2002; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). In many popular reinforcement learning models these prediction errors are used to update the expected value for the option that was chosen on each trial (e.g., Sutton & Barto, 1998; Worthy et al., 2007; Yechiam et al., 2005). These expected values are essentially running averages of the immediate rewards associated with each option, with more weight given to the rewards received most recently. The ability of subcortical regions to track these expected reward values suggests that people have the ability to implicitly learn which option provides the highest immediate payoff on each trial and that this ability may remain intact under WM load (e.g., Frank & Claus, 2006; Frank, O'Reilly, & Curran, 2006).

Although the work cited above suggests that expected reward values based on the immediate rewards received are tracked by subcortical regions that are also associated with implicit, or procedural, learning, it is nevertheless very likely that expected reward values can be explicitly verbalized (e.g., "Option A seems to give larger immediate rewards than Option B"). Thus, it is difficult to distinguish whether the dual WM task leads participants to make decisions based on a more implicit processing mode or whether the dual task leads participants to use a simpler strategy that is overseen by a weakened explicit processing system.

Immediate payoffs are important to consider, but consideration must also be given to each option's delayed payoffs. Often, the consequences of certain decisions are not realized immediately but become manifest at some point in the future. This is similar to many real-world decisions, such as deciding whether to attend graduate school or join the workforce immediately after graduating from college. Joining the workforce may lead to more immediate benefits in the form of higher income than would be earned while in graduate school, but a graduate degree could very likely mean higher cumulative income over the course of one's life. Thus, a second consideration that must be kept in mind when making decisions is how current choices will influence future outcomes or possibilities. People must not be temporally myopic in that they fail to see the delayed consequences of each option as well as the immediate consequences.

A final component of the decision-making process is to efficiently juxtapose the immediate and delayed benefits of each option in order to make the best decision. Although it is very often the case that making decisions based on how those decisions will be rewarded in the future is the best strategy, there are also situations where the immediate benefits outweigh any potential delayed benefits. For example, someone who is nearing retirement age may have little to gain by quitting a job to attend graduate school; similarly, someone may be faced with such a great immediate job prospect that going to graduate or law school may be counterproductive. In some situations the immediate benefits of one option are so valuable that they outweigh any potential de-

layed benefits of other options (Otto, Markman, & Love, 2012; Worthy, Gorlick, Pacheco, Schnyer, & Maddox, 2011).

Our goal in the present work is to investigate the role of WM load in these three common components of decision making:

1. Evaluation of the immediate benefits of each option.
2. Evaluation of the delayed benefits of each option.
3. Juxtaposition of the immediate and delayed benefits to maximize cumulative reward.

One hypothesis is that the amount of necessary WM resources will increase from Components 1 to 3, with the evaluation of the immediate benefits requiring the fewest WM resources and the juxtaposition of the immediate and delayed benefits requiring the most WM resources. To address this issue we use a dynamic decision-making paradigm that is ideal for examining the role of WM load in these three component processes of decision making.

Figure 1 shows the reward structures we use in Experiment 1. We use similar reward structures in Experiment 2, but we vary some aspects of the task to broaden the scope of our findings. In particular, we use different surface features in Experiment 2 that are more similar to gambling paradigms that have been extensively used to examine experience-based decision making (e.g., Bechara et al., 1994; Worthy et al., 2007, 2008; Yechiam et al., 2005). In these tasks there are two options that participants choose from and receive rewards from on each trial. One option, the increasing option, gives a smaller reward on each trial, but selecting this option causes rewards for both options to increase on future trials. This can be seen on the *x*-axes for Figures 1a and 1b, where rewards are a direct function of the number of times the increasing option has been selected over the previous 10 trials. The other option, the decreasing option, gives a larger reward on any given trial, but selecting this option causes future rewards for both options to decrease. Thus, selecting the decreasing option always leads to a larger immediate reward, and selecting the increasing option always leads to larger rewards for both options on future trials. Tendencies to select either of these options can indicate a preference or bias toward immediate versus delayed payoffs, and we can thus address how WM load influences these preferences.

To address how WM load affects the third component of decision making listed above, the juxtaposition of the immediate and delayed benefits of each option, we alter which option is optimal between the reward structures in Figure 1a and 1b. Figure 1a shows the reward structure for Experiment 1's *increasing-optimal* task. Here the increasing option is the optimal choice because repeatedly selecting it will lead to more oxygen earned on each trial (80 points) than the amount earned from repeatedly selecting the decreasing option (40 points). In contrast, in the *decreasing-optimal* task shown in Figure 1b, the decreasing option gives a much larger reward on each trial than the increasing option (60 points more). Here the amount earned from repeatedly selecting the increasing option (55 points) is smaller than the amount earned from repeatedly selecting the decreasing option (65 points), so selecting the increasing option is futile and counterproductive even though it leads to more rewards for both options on future trials. To perform well on both of these tasks, participants must not simply be biased toward the immediate versus delayed benefits of the two

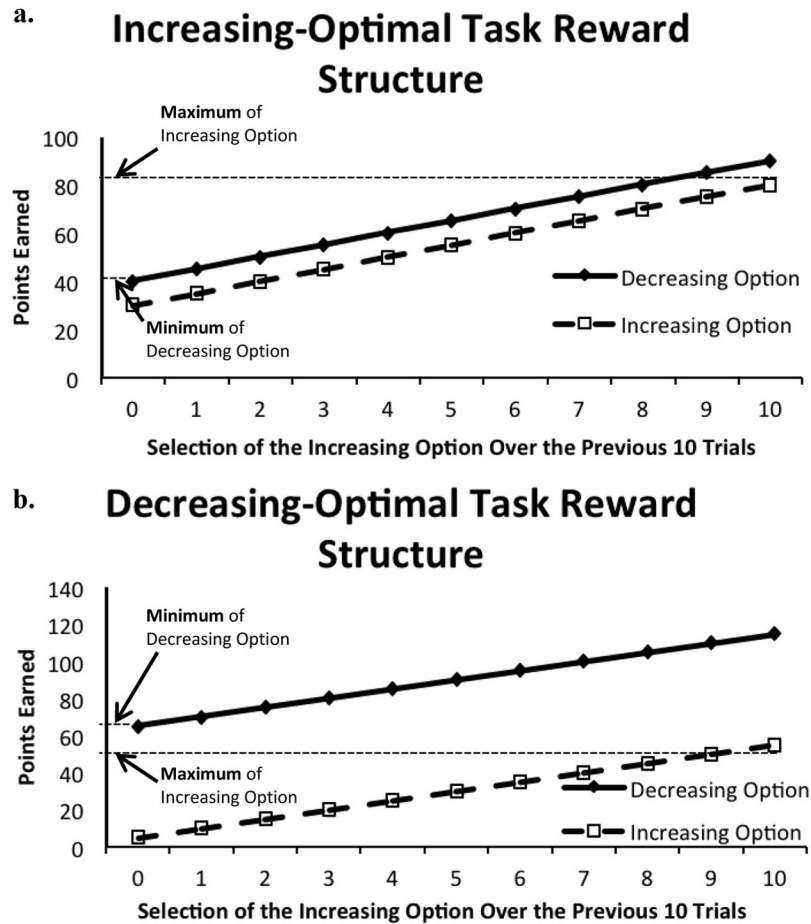


Figure 1. Rewards given for each task as a function of the number of increasing option selections over the previous 10 trials. (a) Rewards given for the increasing-optimal task. Selecting the increasing option 10 consecutive times will lead to a reward of 80 points on each trial, whereas selecting the decreasing option 10 consecutive times will lead to a reward of 40 points on each trial. (b) Rewards given for the decreasing-optimal task. Selecting the decreasing option 10 consecutive times will lead to a reward of 65 points on each trial, whereas selecting the Increasing option 10 consecutive times will lead to a reward of 55 points on each trial.

options but instead must effectively juxtapose the benefits of the two options.

We examine the role of WM load by having participants make decisions with or without performing a concurrent numerical Stroop task that is designed to deplete available WM resources. This task has been successfully used in previous published work to examine the role of WM load during category learning (Miles & Minda, 2011; Newell, Dunn, & Kalish, 2010; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). This research has found that depleting WM resources through the use of the numerical Stroop task results in impaired performance on explicit, rule-based tasks, where a verbalizable rule can distinguish members of each category, but it does not impair performance on procedural, information-integration tasks, where verbalizable rule use is difficult (but see Newell et al., 2010, and Nosofsky & Kruschke, 2002, for a counterargument, and Ashby & Ell's 2002 reply to Nosofsky & Kruschke, 2002). The numerical Stroop task requires participants to maintain information about two different properties of the numbers presented in working memory while making each deci-

sion (the numerical value of the number as well as the font size used to display the number).

Predictions

We offer two possible hypotheses for how performing under dual-task versus single-task conditions will affect dynamic decision making. One possibility is that participants performing under dual-task conditions will simply underperform on both increasing-optimal and decreasing-optimal tasks. Previous research has demonstrated poor choice performance under WM load for participants performing the Iowa Gambling task (Dretsch, & Tipples, 2008; Hinson et al., 2002), and reduced WM resources from the dual task may simply attenuate participants' ability to appropriately weigh the immediate and delayed benefits of each option. Additionally, some previous research suggests that WM load may simply lead to more random responding (Franco-Watkins, Pashler, & Rickard, 2006). However, this conclusion has been challenged by other

findings that suggest that participants under WM load are not simply behaving randomly (e.g., Dretsch & Tipples, 2008).

A second possibility, which we view as more likely, is that the dual task will cause a type of temporal myopia so that participants under WM load are unable to observe the long-term effects of selecting each option. This may result from the WM load causing dual-task participants to learn the expected values for each option more implicitly based on the immediate feedback received after each selection and the resultant prediction error. Some previous work suggests that WM load may actually increase implicit forms of learning due to a reduction in interference from the explicit system when the task requires an implicit strategy (e.g., Filoteo, Lauritzen, & Maddox, 2010; Maddox, Ashby, Ing, & Pickering, 2004). Single-task participants may rely less on immediate feedback and may use a heuristic-based strategy that focuses on comparing the rewards received across trials. Such a comparison would allow single-task participants to observe the delayed effects of choosing each option. One possible strategy that single-task participants may employ is a win-stay, lose-shift (WSLS) strategy (Nowak & Sigmund, 1993; Otto, Taylor, & Markman, 2011; Steyvers, Lee, & Wagenmakers, 2009). This is a rule-based strategy that has been shown to be used in binary prediction tasks (e.g., Otto et al., 2011). Under this strategy, participants “stay” by picking the same option on the next trial if they were rewarded and “switch” by picking the other option on the next trial if they were not rewarded.

Here, we examine whether participants use a more elaborate version of this strategy in the present tasks by examining whether participants “stay” by picking the same option on the next trial if the reward was equal to or larger than reward received on the previous trial (a “win” trial) or “shift” by selecting the other option on the next trial if the reward received on the current trial was smaller than the reward received on the previous trial (a “lose” trial). After examining the behavioral data to determine if there are differences in the proportion of times single- and dual-task participants select each option and how well they perform on the tasks, we then fit the data with two learning models: a Softmax reinforcement learning model, which assumes that expected reward values are updated each time an option is chosen based on the rewards received immediately after each selection, and a WSLS model, which assumes that participants adjust their behavior based on a comparison between the current reward and the reward received on the previous trial. These models both have long histories in the decision-making literature. The former model may mimic a more implicit or procedural decision-making mode, and the latter model likely mimics a more explicit, or rule-based, decision-making mode. A WSLS strategy also requires participants to remember the reward that was received on the previous trial. The dual task will make doing so more difficult. Accordingly, we predict that behavior of single-task participants will be better described by the WSLS model and that behavior of dual-task participants will be better described by the Softmax reinforcement learning model. The use of these different strategies should lead single-task participants to select the increasing option more often than dual-task participants do across all conditions. In the next section we present the details of the WSLS and Softmax models and then present simulations of the tasks shown in Figure 1 for each model. These simulations provide predictions for how par-

ticipants will perform if they are using WSLS or Softmax-based strategies.

Predictions From the WSLS and Softmax Models

The WSLS and Softmax models have been used in many previous studies to characterize decision-making behavior, and the assumptions and mechanisms of each are transparent (Frank & Kong, 2008; Lee, Zhang, Munro, & Steyvers, 2011; Otto et al., 2011; Steyvers et al., 2009; Sutton & Barto, 1998; Worthy & Maddox, 2012; Worthy et al., 2007).

The Softmax model assumes that participants develop expected values (EVs) for each option that represent the rewards they expect to receive upon selecting each option. EVs for all options are initialized at 0 at the beginning of the task and are updated only for the chosen option, i , according to the following updating rule:

$$EV_{i,t+1} = EV_{i,t} + \alpha \cdot [r(t) - EV_{i,t}] \quad (1)$$

Learning is modulated by a learning rate, or recency, parameter (α), $0 \leq \alpha \leq 1$, that weighs the degree to which the model updates the EVs for each option based on the prediction error between the reward received [$r(t)$], and the current EV on trial t . As α approaches 1, greater weight is given to the most recent rewards in updating EVs, indicative of more active updating of EVs on each trial, and as α approaches 0, rewards are given less weight in updating EVs. When $\alpha = 0$ no learning takes place, and EVs are not updated throughout the experiment from their initial starting points.

The EVs for all j options are used to determine the model’s probability for selecting each option. Action selection probabilities for each option (a) are computed via a Softmax decision rule (Sutton & Barto, 1998):

$$P(a_i) = \frac{e^{\{\gamma \cdot EV_{a,i}\}}}{\sum_{j=1}^4 e^{\{\gamma \cdot EV_{j,i}\}}} \quad (2)$$

Here γ is an exploitation parameter that determines the degree to which the option with the highest EV is chosen. As γ approaches infinity the highest valued option is chosen more often, and as γ approaches 0 all options are chosen equally often. This model has been used in a number of previous studies to characterize choice behavior (e.g., Daw, O’Doherty, Seymour, & Dolan, 2006; Otto et al., 2010; Worthy et al., 2007; Yechiam & Bussemeyer, 2005).

The WSLS model has two free parameters: the probability of staying with the same option on the next trial if the reward received on the current trial is equal to or greater than the reward received on the previous trial, $P(stay|win)$, and the probability of shifting to the other option on the next trial if the reward received on the current trial is less than the reward received on the previous trial, $P(shift|loss)$.

To predict how participants would perform if they were relying on a WSLS versus a Softmax-based strategy we conducted 1,000 simulations for each model for each of the two tasks shown in Figure 1. We selected reasonable parameter values to use in the simulations based on parameter values for each model that provided a good fit to data from published work in our labs (Worthy et al., 2011). The parameters used for the WSLS model were .90

for $P(stay|win)$ and .5 for $P(shift|loss)$, and the parameters used for the Softmax model were .50 for α and .03 for γ .

Figure 2 shows the average proportion of trials that each model selected the increasing option in each task. The WSLS model selected the increasing option more than the Softmax model did in both tasks, and there was no difference between the two tasks. This is because the model’s decisions are based only on whether the current reward is greater or less than the reward given on the previous trial, regardless of the magnitude of the difference. The Softmax model, which is sensitive to the magnitude of the difference in average rewards given for each option, selected the increasing option less often than the WSLS model did in both tasks but much less often in the decreasing-optimal task, where the decreasing option gives a substantially higher reward than the increasing option on each trial. We predict, based on these simulations, that single-task participants, who are predicted to rely on a WSLS strategy more than dual-task participants, will select the increasing option more than dual-task participants will in both tasks. This will lead to better performance in the increasing-optimal task but worse performance than dual-task participants in the decreasing-optimal task.

Experiment 1

Method

Participants. Ninety-eight undergraduate students at Texas A&M University participated in the experiment for course credit. Participants were randomly assigned to one of four between-subjects conditions that consisted of the factorial combination of two WM load conditions (single-task vs. dual-task) and two reward structure conditions (increasing-optimal vs. decreasing-optimal). There were 23 participants in the dual-task decreasing-optimal condition and 25 participants in each of the other three conditions.

Materials and procedure. Participants performed the experiment on PCs using Psychtoolbox for Matlab (version 2.5).

The procedure for the single-task conditions was nearly identical to the procedure used in a previous study in our labs (Worthy et al., 2011). Figure 3 shows a sample screenshot from the experiment. Participants were given a cover story that they would be

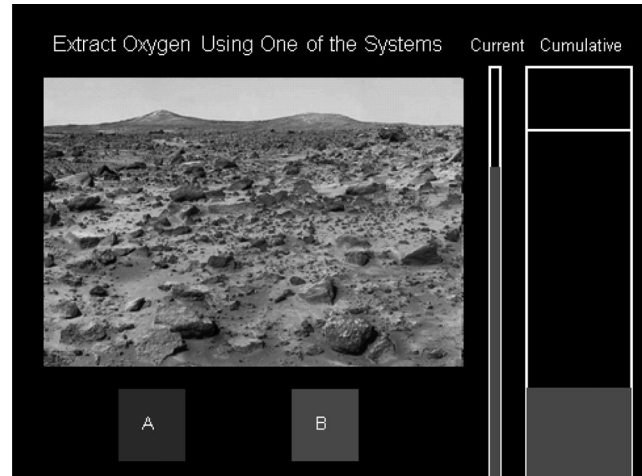


Figure 3. Sample screen shot from the experiment. Participants were given a cover story in which they were asked to test two oxygen-extraction systems on the Martian landscape. The oxygen extracted on each trial was shown in the Current tank and then transferred to the Cumulative tank.

testing two extraction systems that farmed oxygen on Mars, and their goal was to extract as much oxygen as possible. A similar paradigm has been used elsewhere to examine other issues in dynamic decision making (Gureckis & Love, 2009a, 2009b; Otto et al., 2009; Otto et al., 2012). On each trial participants were told to “collect oxygen using one of the two systems” that appeared at the top of the screen. They were allowed as much time as they wished to make a response. After a delay of 500 ms the amount of oxygen they received for that trial was indicated in the narrow tank labeled “Current,” and after another 1,000 ms the oxygen would be emptied into the “Cumulative” tank. To roughly equate the experiment time between WM load conditions, there was a 1,000-ms intertrial interval in the single-task conditions in which a black screen came up and participants saw the phrase, “Please wait for the next trial . . .” The next trial would then begin.

Participants performed a total of 250 trials of the task. The rewards they received were based on the reward structures shown in Figure 1. Participants in the increasing-optimal condition received rewards plotted in Figure 1a, and participants in the decreasing-optimal condition received rewards plotted in Figure 1b. The place on the x-axis was determined on each trial by summing the number of times participants had selected the increasing option over the previous trial. Thus, there was a “moving window” that kept a count of the number of times the increasing option was selected on each trial. All participants began the experiment at the midpoint (5) on the x-axis.

A line on the larger tank corresponded to the amount of oxygen needed to sustain life on Mars. Participants were given the goal of trying to collect this amount of oxygen over the course of the experiment. The goal lines were set at the equivalent of 18,000 points for the increasing-optimal task and 16,000 points for the decreasing-optimal task. This corresponded to selecting the optimal choice in each task on roughly 80% of trials. Participants were told nothing about the rewards available for each option or the choice history–dependent structure of the rewards.

The procedure for participants in the dual-task conditions was modified from the procedure for the single-task conditions to

Increasing Option Selections from Simulations of Each Model

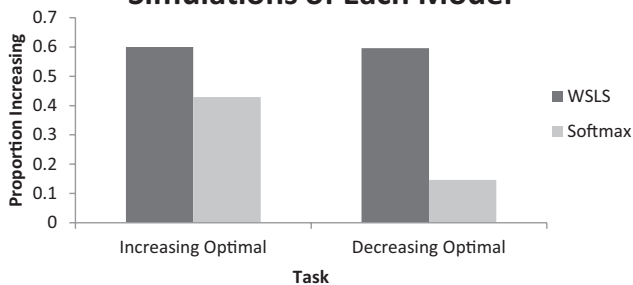


Figure 2. Average proportion of increasing options selected from each model’s simulations. WSLS = win-stay, lose-shift model; Softmax = Softmax reinforcement learning model.

include a numerical Stroop task used in previous work to investigate the role of WM load in perceptual category learning (Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). Participants had to perform a numerical analogue of the Stroop task (Stroop, 1935) concurrently with the decision-making task. The concurrent task required participants to remember which of two numbers was physically larger and which was larger in numerical value and to hold that information in mind while deciding which oxygen system to choose on each trial. On each trial the picture of Mars, the two extraction systems, and the Current and Cumulative point meters were presented in the center of the screen. At the beginning of the trial the two numbers for the concurrent task were presented on each side of the screen, one number on each side, for 200 ms. After 200 ms a uniform white mask covered the numbers for 200 ms. Participants were then allowed to choose from one of the two oxygen extraction systems, and they were given feedback in a manner identical to that described above. A new screen then appeared that queried participants with either *VALUE* or *SIZE*, and participants selected from buttons labeled *Left* or *Right* to indicate which side had the number largest in either numerical value or physical size. After they made their response they were told whether they were correct or not, and the next trial began immediately.

Following previous studies that have used the same concurrent task manipulation, participants were told that they were required to achieve at least 80% accuracy on the numerical Stroop task and that they should focus first on achieving good performance on the numerical Stroop task and “use what you have left over” for the decision-making task (e.g., Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). Participants’ current accuracy on the numerical Stroop task was indicated at the top of the screen when they received feedback regarding their performance on the concurrent task on each trial. Their percentage correct score was listed in green if it was above 80% and in red if it was below 80%. All of the participants in the study achieved at least 80% accuracy on the numerical Stroop task.²

Results

Behavioral analyses. The effects of WM load on behavior can be seen by simply plotting the proportion of times participants selected the increasing option over the course of the experiment (see Figure 4a). A 2 (WM load) \times 2 (reward structure) analysis of variance (ANOVA) revealed a significant effect of WM load, $F(1, 94) = 17.46, p < .001, \eta^2 = .16$. Participants in the single-task ($M = .49, SD = .21$) condition selected the increasing option much more often than participants did in the dual-task condition ($M = .30, SD = .24$). There was also a significant effect of task type, $F(1, 94) = 7.73, p < .01, \eta^2 = .08$. Participants who performed the increasing-optimal task ($M = .45, SD = .24$) selected the increasing option more often than participants who performed the decreasing-optimal task did ($M = .33, SD = .23$). The WM Load \times Task Type interaction was not significant ($F < 1$).

To compare performance across tasks we derived a measure of the proportion of the optimal cumulative payoff that is commensurable across task structures. This measure was derived by computing (points earned – minimum possible points)/range of possible points earned. Because points received are a function of the proportion of times a participant selects each option, the proportion

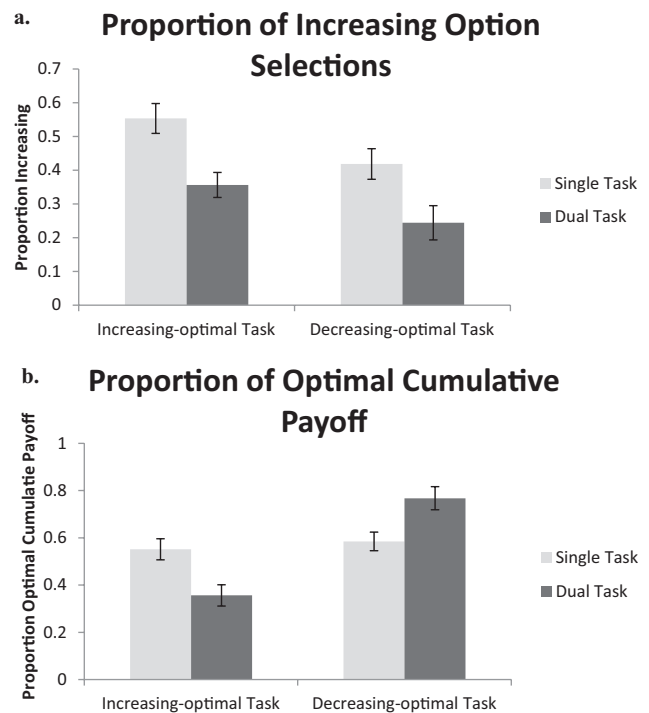


Figure 4. (a) Proportion of increasing options selections for participants in each condition of Experiment 2. (b) Proportion of the optimal cumulative payoff participants earned in the task. These proportions are roughly equivalent to the proportion of times participants selected the optimal option based on the reward structure. Error bars represent standard errors.

of the optimal cumulative payoff value is an indirect measure of the proportion of times participants made the optimal choice for each reward structure. We used this measure rather than points earned on the task, because it equated performance across the two tasks, which differed in their reward structures. Figure 3b plots the proportions of the optimal cumulative payoff for participants in each condition. A 2 (WM load) \times 2 (reward structure) ANOVA revealed a significant effect of reward structure, $F(1, 94) = 25.11, p < .001, \eta^2 = .21$. Participants in the decreasing-optimal reward structure conditions ($M = .67, SD = .23$) earned higher proportions of the optimal cumulative payoff than participants in the increasing-optimal conditions did ($M = .45, SD = .24$). There was also a significant WM Load \times Reward Structure interaction, $F(1, 94) = 18.15, p < .001, \eta^2 = .16$. We conducted pairwise comparisons within each reward structure to examine the locus of this interaction. For the increasing-optimal reward structure, single-task participants ($M = .55, SD = .22$) significantly outperformed dual-task participants ($M = .36, SD = .22$), $F(1, 48) = 9.60, p < .01, \eta^2 = .17$. However, for the decreasing-optimal reward structure, dual-task participants ($M = .77, SD = .23$) significantly outperformed single-task participants ($M = .58, SD = .20$), $F(1, 46) = 8.59, p < .01, \eta^2 = .16$.

² We did not observe any correlations with performance on the WM task and performance in the decision-making task in either experiment. Most participants performed well above 80%, and so there was little variance in performance on the WM task within the dual-task condition.

We also examined the mean response times for participants in each condition. These are shown in Figure 5. A 2 (WM load) \times 2 (reward structure) ANOVA revealed a significant effect of WM load, $F(1, 94) = 25.11, p < .001, \eta^2 = .21$. Dual-task participants ($M = 1.44, SD = 0.62$) took significantly longer to make a decision than single-task participants ($M = 0.79, SD = 0.38$). One possibility for the longer response times for dual-task participants is that they spent time rehearsing the information for the numerical Stroop task before making a decision on each trial.

Model-based analyses. We fit each participant's data individually with the Softmax and WLSL models. We also fit a baseline or null model that assumes fixed choice probabilities (Gureckis & Love, 2009a; Worthy & Maddox, 2012; Yechiam et al., 2005). This model has one free parameter, $p(\text{Increasing})$, that represents the probability of selecting the increasing option on any given trial. The probability of selecting the decreasing option is $1 - p(\text{Increasing})$. Although this model does not assume that participants learn from rewards given on each trial, it can often provide a good fit to the data (Gureckis & Love, 2009a). Indeed, the optimal strategy in each task is to select the best option repeatedly, and data from a participant who exhibited such maximizing behavior would be well fit by the baseline model. Thus, a good fit for the baseline model does not necessarily imply random responding.

The models were assessed based their ability to predict each choice a participant would make on the next trial, by estimating parameter values that maximized the log-likelihood of each model given the participant's choices. We used Akaike's information criterion (AIC) (Akaike, 1974) to examine the fit of the WLSL and Softmax models relative to the fit of the baseline model. AIC penalizes models with more free parameters. For each model, i , AIC_i is defined as

$$AIC_i = -2\log L_i + 2V_i \quad (3)$$

where L_i is the maximum likelihood for model i , and V_i is the number of free parameters in the model. Smaller AIC values indicate a better fit to the data. We compared the fits of the WLSL and Softmax models relative to fits of the baseline model by subtracting the AIC of each model from the AIC of the baseline model for each participant's data (e.g., Gureckis & Love, 2009a):

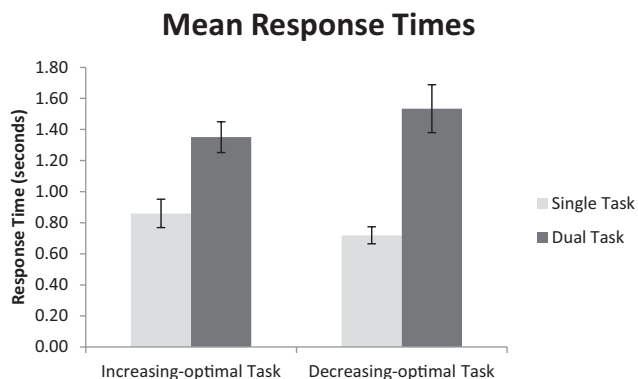


Figure 5. Average response times for participants in each condition. Error bars represent standard errors.

$$Relative\ Fit(M) = AIC_B - AIC_M \quad (4)$$

Positive values indicate a better fit of the learning model, and negative values indicate a better fit of the baseline model.

The relative fits for each model are listed in Table 1. Data from single-task participants are better fit by the WLSL model, and those from dual-task participants are better fit by the Softmax model. Overall, relative fit values were lower for dual-task participants, although data from single-task participants who performed the decreasing-optimal task were fit no better by the Softmax model than the baseline model. This could suggest that the behavior of dual-task participants is more random than the behavior of single-task participants (cf. Franco-Watkins et al., 2006), or that they followed a near-deterministic response process by selecting the decreasing option on the majority of trials. However, average relative fits of the Softmax model for dual-task participants were well above 0 in both tasks, which suggests that this model provided a good fit to a large number of dual-task participants' data. We also examined the proportion of data sets that were fit best by the baseline model. The baseline model provided the best fit for 20% of dual-task participants in the increasing-optimal task and 39% of participants in the decreasing-optimal task, compared to only 4% and 12% of single-task participants in the increasing and decreasing-optimal tasks, respectively. These differences in the proportion of single- and dual-task participants that were best fit by the baseline model are both significant ($p < .05$ by binomial test). Thus, there was more evidence for dual-task participants relative to single-task participants being best fit by the baseline model, but there was also much evidence that dual-task participants were fit well by the Softmax model. This suggests that they were not simply behaving randomly.

Having established that participants were by and large well characterized by models that exhibit dependence on previous actions and outcomes rather than a null model, we next compared the relative fit of the WLSL model and the Softmax model. To obtain a relative measure of the degree to which the WLSL model provided a better fit to the data, we subtracted the log-likelihood of the WLSL model from the log-likelihood of the Softmax model ($Relative\ fit_{WLSL} = \ln L_{Softmax} - \ln L_{WLSL}$). Because these models have the same number of free parameters, their log-likelihood values can be compared directly. Under our likelihood ratio metric, positive $Relative\ fit_{WLSL}$ values indicate a better fit for the WLSL model, and negative $Relative\ fit_{WLSL}$ values indicate a better fit for the Softmax model.

Figure 6 shows the $Relative\ fit_{WLSL}$ values for participants in each condition. We performed nonparametric Mann-Whitney U tests to examine the effects of WM load and reward structure, as the $Relative\ fit_{WLSL}$ distributions were markedly nonnormal. There was a significant effect of WM load ($U = 564.50, p < .001$). Single-task participants ($M = 45.35, SD = 66.36$) had much higher $Relative\ fit_{WLSL}$ values than dual-task participants did ($M = -11.63, SD = 22.40$). This suggests that single-task participants were more likely to use a WLSL strategy and that dual-task participants were more likely to use a Softmax strategy. There was no significant effect of reward structure ($U = 1,316, p > .10$).

Parameter estimates. The average estimated parameter values for participants in each condition are listed in Table 2. We conducted a 2 (WM load) \times 2 (reward structure) ANOVA on the average estimated recency parameter (α) values for participants in

Table 1
Relative Fits of the WLS and Softmax Models to the Baseline Model for Experiment 1

Task and condition	WLS	Softmax
Increasing-optimal task		
Single-task	60.15 (12.56)	29.51 (9.41)
Dual-task	4.32 (8.29)	16.93 (6.99)
Decreasing-optimal task		
Single-task	59.79 (14.23)	-0.20 (3.57)
Dual-task	3.18 (7.57)	13.75 (6.38)

Note. Positive values indicate a better fit of each model than for the baseline model. Values shown are means and standard errors of the mean (SE). WLS = win-stay, lose-shift; Softmax = Softmax reinforcement learning.

each condition. There was a marginally significant effect of WM load, $F(1, 94) = 3.11, p < .10, \eta^2 = .03$. Data from participants in the single-task condition were better fit by higher recency parameter values ($M = .39, SD = .45$) than data from participants in the dual-task condition were ($M = .24, SD = .36$). There was no effect of reward structure and no WM Load \times Reward Structure interaction ($F_s < 1$). The difference in recency parameter values between single- and dual-task participants is consistent with the results of Otto et al. (2011), who found higher estimated recency parameter values for single-task participants who performed a simple binary prediction task. Otto et al. also found increased reliance on a WLS strategy among single-task participants compared to dual-task participants. One interpretation of this difference is that single-task participants are more responsive to recent outcomes. One way in which the Softmax model may account for data from participants who are using a WLS strategy is by better fitting the data with higher estimated recency parameter values. This suggests that there may be an association between recency parameter values and $Relative\ fit_{WLS}$. We did indeed find such a correlation ($r = .23, p < .05$).³

A 2 (WM load) \times 2 (reward structure) ANOVA on the average estimated exploitation parameter (γ) values for participants in each condition revealed no effect of WM load or reward structure and no interaction ($F_s < 1$). For the $P(stay|win)$ parameter, a 2 (WM load) \times 2 (reward structure) ANOVA showed no effect of WM load, $F(1, 94) = 1.28, p > .10$, and no effect of reward structure, $F(1, 94) = 1.08, p > .10$. However, there was a significant WM Load \times Reward Structure interaction, $F(1, 94) = 4.12, p < .05, \eta^2 = .04$. For the increasing-optimal task, data from single-task participants ($M = .83, SD = .09$) were best fit by $P(stay|win)$ parameter values significantly higher than those for data from dual-task participants ($M = .74, SD = .17, F(1, 94) = 5.52, p < .05, \eta^2 = .10$). However, there was no difference between single- and dual-task participants for the decreasing-optimal task ($F < 1$).

For the $P(shift|loss)$ parameter, a 2 (WM load) \times 2 (reward structure) ANOVA revealed a main effect of WM load, $F(1, 94) = 8.55, p < .01, \eta^2 = .08$. Data from single-task participants ($M = .37, SD = .18$) were best fit by $P(shift|loss)$ parameter values lower than those for data from dual-task participants ($M = .46, SD = .12$). There was also a main effect of reward structure, $F(1, 94) = 9.95, p < .01, \eta^2 = .10$. Estimated $P(shift|loss)$ parameters were higher for data from participants who performed the increasing-

optimal task ($M = .46, SD = .15$) than for data from participants who performed the decreasing-optimal task ($M = .37, SD = .15$). The WM Load \times Reward Structure interaction was not significant.

Cross-fitting analysis. Recent work has demonstrated that model complexity cannot be solely accounted for by measures like AIC or the Bayesian information criterion (BIC: Schwarz, 1978) to penalize models for the number of free parameters (Djuric, 1998; Myung & Pitt, 1997). Often, models with the same number of free parameters can differ in how flexible they are in accounting for data. For example, the WLS model may be more flexible than the Softmax model in that it can account for a wider range of behavior in decision-making tasks, or the Softmax model may be more flexible in that it can account for data from participants who are using a WLS strategy with higher recency parameter values. To address this issue we used a procedure known as the parametric bootstrap cross-fitting method (PBCF) proposed by Wagenmakers and colleagues (Wagenmakers, Ratcliff, Gomez, & Iverson, 2004; see also Donkin, Brown, Heathcote, and Wagenmakers, 2011, for a similar approach). This method involves simulating a large number of data sets with each of two models and then fitting each data set with each model. If neither model can mimic the other model, the model that generated the data should provide the best fit to the majority of data sets. However, if a large proportion of data sets that are generated by one model is fit better by the non-data-generating model, that would suggest that the non-data-generating model is more flexible in its ability to mimic the true data-generating model.

We used sets of parameter values that were estimated from our participants' data for the simulated data sets. For each task we generated 2,000 data sets using parameter combinations that were sampled with replacement from the best fitting parameter distributions for participants in each experiment. Thus, for the Softmax model we randomly sampled a combination of α and γ that provided the best fit to one participant's data and used those parameter values to perform one simulation of the task. We generated 2,000 simulated data sets in this manner and performed the same simulation procedure with the WLS model. We then fit each simulated data set with both models and determined the $Relative\ fit_{WLS}$ value for each data set as outlined above.

Figure 7 plots histograms of the $Relative\ fit_{WLS}$ values for data generated by each model for the increasing-optimal task (see Figure 7a) and the decreasing-optimal task (see Figure 7b). These values should be less than zero for data generated by the Softmax model and greater than zero for data generated by the WLS model if neither model is able to mimic the other model. For the increasing-optimal task, 85% of data sets generated by the Softmax model were also fit best by the Softmax model, and 98% percent of data sets generated by the WLS model were also fit best by the WLS model. This suggests that the WLS model has a higher probability of mimicking the Softmax model than vice versa. The overall probability of correct classification of the data-generating model is $(.85 + .98)/2 = .915$ (Wagenmakers et al., 2004).

For the decreasing-optimal task, 71% of data sets generated by the Softmax model were also fit best by the Softmax model, and

³ A reanalysis of the data reported by Otto et al. (2011) found a strong relationship between estimated recency parameter values and the relative fit of the WLS model used to fit their data ($r = .87, p < .001$).

Fit of the WSL Model Relative to the Softmax Model

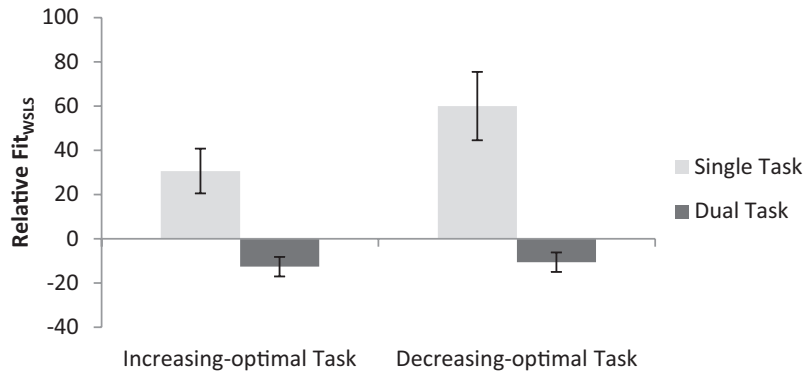


Figure 6. Average relative fit values of the WSL model compared to the Softmax reinforcement learning model for data from Experiment 1. Higher values indicate a better fit for the WSL model. Error bars represent standard errors. WSL = win-stay, lose-shift; Softmax = Softmax reinforcement learning.

97% of data sets generated by the WSL model were also fit best by the WSL model. The overall probability of correct classification of the data generating model is .84. These results suggest (a) that there is a high probability of correctly classifying the model as the true data-generating model and the probability is higher for the increasing-optimal task than the decreasing-optimal task (.915 vs. .840) and (b) that the WSL model is better able to mimic the Softmax model than vice versa.

However, it should be noted that the majority of data sets that were generated by the Softmax model but better fit by the WSL model had $Relative\ fit_{WSL}$ values that were only slightly greater than 0, indicating only a slightly better fit by the WSL model. This is indicated by the large spikes around the zero point in the x -axis in Figures 7a and 7b. Thus, for many data sets the models provided essentially the same fit to the data. The average $Relative\ fit_{WSL}$ values for single-task participants in both tasks were extremely high ($M = 30.65$ for the increasing-optimal task and $M = 59.99$ for the decreasing-optimal task). Our cross-fitting analysis suggests that it is unlikely that these large $Relative\ fit_{WSL}$ values for single-task participants resulted from the WSL model mimicking the Softmax model. Thus, although the WSL model may be able to mimic the Softmax model to some extent, it is highly unlikely that single-task participants who were best fit by the

WSL model were actually using a strategy that is consistent with the Softmax model.

Discussion

The data clearly indicate that WM load modulates the degree to which participants attend to the immediate versus delayed consequences of their actions, and the model-based analyses support this conclusion. We used an experimental paradigm that allowed us to examine this immediate versus delayed dichotomy by offering two choices: one that always led to a larger immediate reward (the decreasing option) and one that always led to larger rewards for both options on future trials (the increasing option), despite giving a smaller immediate payoff on each trial. The paradigm also allowed us to manipulate which of the two options was the optimal choice for the task, and this allowed us to examine whether WM load affected the preference for the immediate versus delayed rewarding option or whether it affected participants' ability to juxtapose the benefits of each option. Both our behavioral and model-based analyses clearly show that WM load affected only the preference for immediate versus delayed rewarding option, with participants in the single-task conditions preferring the increasing option, which led to larger delayed rewards, in both tasks more than participants in the dual-task conditions did. Thus, the WM load brought about by the concurrent dual task hurt in one case but helped in the other.

The model-based analyses show a clear effect of WM load on the relative fit of two models that assume different decision-making strategies. Behavior of single-task participants was best fit by the WSL model, which assumes that participants maintain the reward received on the last trial in WM and use it as a benchmark for deciding whether to stay with the option that was picked or shift to the other option. A WSL strategy will allow participants to notice how each option causes future rewards to either increase or decrease. For example, if participants repeatedly select the increasing option they will repeatedly "win," and if participants repeatedly select the decreasing option they will repeatedly "lose."

Table 2
Average Best Fitting Parameter Values for Participants in Each Condition in Experiment 1

Task and condition	$P(stay win)$	$P(shift loss)$	α	γ
Increasing-optimal task				
Single-task	.82 (.02)	.41 (.04)	.36 (.09)	.12 (.04)
Dual-task	.73 (.03)	.51 (.02)	.21 (.07)	.19 (.06)
Decreasing-optimal task				
Single-task	.80 (.02)	.33 (.03)	.43 (.09)	.12 (.05)
Dual-task	.82 (.41)	.41 (.03)	.27 (.08)	.14 (.05)

Note. Standard errors of the mean are listed in parentheses.

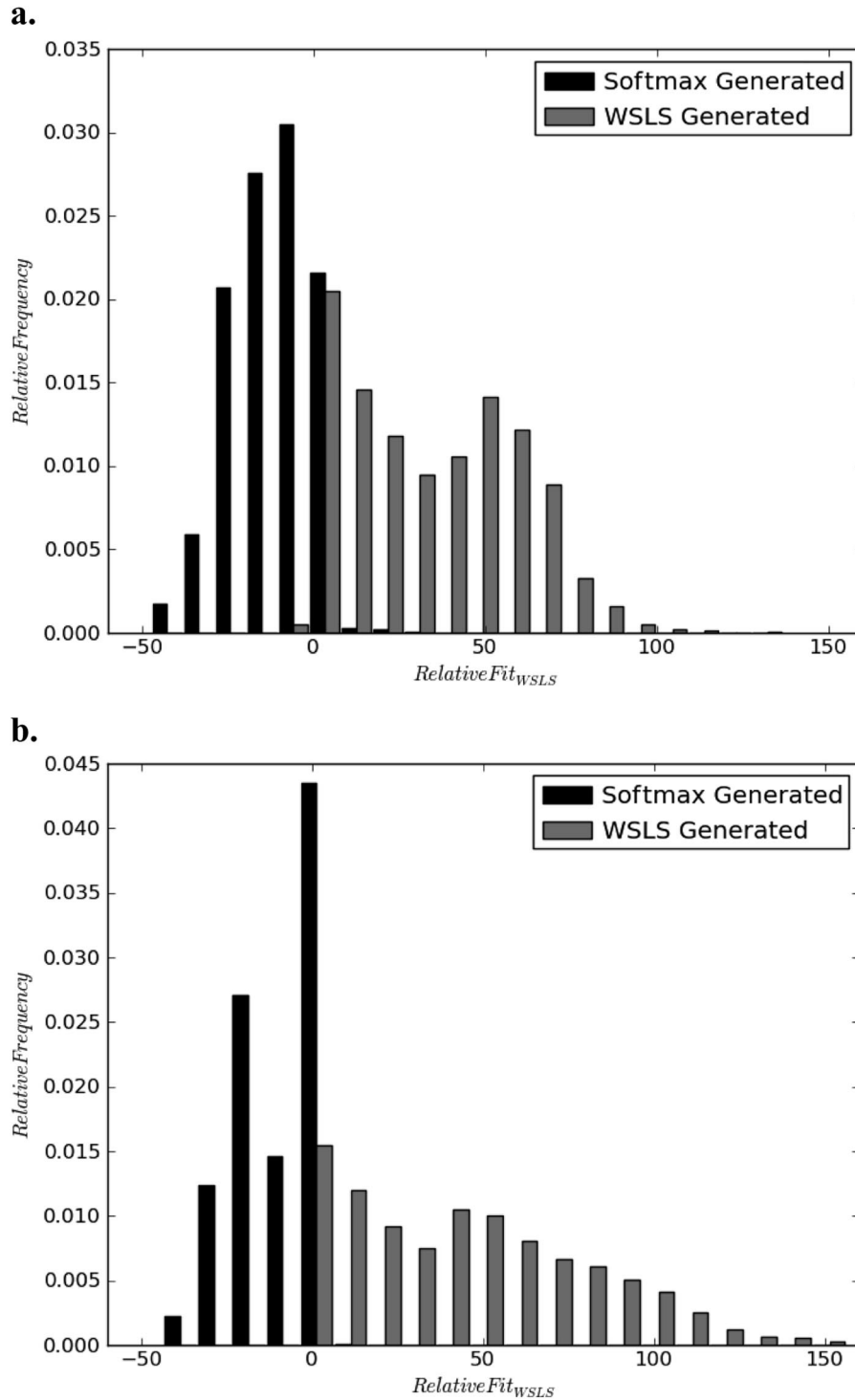


Figure 7. Distributions of fits of the WSLs model relative to the Softmax model for data simulated by each model for the (a) increasing-optimal and (b) decreasing-optimal tasks. Values greater than zero indicate a better fit for the WSLs mode, values less than zero indicate a better fit for the Softmax model, and values equal to zero indicate an equal fit for the two models. WSLs = win-stay, lose-shift; Softmax = Softmax reinforcement learning.

Data from dual-task participants were fit less well by the WSLs model than the Softmax reinforcement learning model. This suggests that WM load led participants to update their expected values for each option based only on the rewards received immediately after each selection. The WM load manipulation may have prevented dual-task participants from noticing how previous choices affected rewards on future trials. The modeling results also suggest that there may be separate implicit and explicit systems that operate in different ways. The implicit system may develop expected values based on immediate feedback and compare the expected values for each option to make decisions; the explicit system may use verbalizable rules, such as “pick the same option if the reward was greater than or equal to the reward received on the previous trial, or switch options if the reward received was less than the reward received on the previous trial,” and makes decisions based on the rule being used (e.g., Otto et al., 2011). This is similar to a prominent dual systems view in category learning that posits the existence of an explicit, rule-based system and a procedural, information-integration system that relies on immediate feedback (e.g., Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Maddox, 2005; Maddox & Ashby, 2004). The results are also consistent with a long line of research that has posited the existence of dual learning systems, with a major distinction being between a WM-dependent, explicit system and a non-WM-dependent, implicit associative system (Evans, 2008; Metcalfe & Mischel, 1999; Poldrack & Packard, 2003; Sloman, 1996; Smith & DeCoster, 2000; Strack & Deutsch, 2004).

Although the two tasks used in Experiment 1 differed in the choice that was optimal, they also differed in how many points would be rewarded if one consistently picked the optimal versus the inferior option. In the increasing-optimal task the participant earned 80 points for selecting the Increasing option repeatedly, whereas only 40 points was earned for selecting the decreasing option repeatedly. In the decreasing-optimal task the participant earned 65 points for selecting the decreasing option repeatedly, whereas only 55 points was earned for selecting the increasing option repeatedly. Thus, what we will call “end-state separation”—the difference in points earned between repeatedly selecting the optimal option versus repeatedly selecting the sub-optimal option—differed between the two tasks, with the increasing-optimal task having a larger end-state separation (40 points) than the decreasing-optimal task (10 points) did.

In Experiment 2 we conceptually replicated the effect of WM load on decision making and equated end-state separation between the increasing-optimal and decreasing-optimal tasks. We had participants perform either the increasing-optimal task from Experiment 1 or a decreasing-optimal task with a large end-state separation (40 points, which is the end-state separation for the increasing-optimal task), under either single- or dual-task conditions. These reward structures are shown in Figure 8.

Here, we used a more traditional gambling task paradigm rather than the Mars farming cover story used in Experiment 1. Figure 9 shows a sample screenshot from the experiment. A main difference between the Mars farming task and gambling task paradigms, of potential relevance to the WM task manipulation, is the form of feedback given after each decision. In the Mars farming task setup, feedback is given in the form of the amount of oxygen shown in the narrow tank after each selection, and no numbers are shown to represent the amount of oxygen given. In the gambling task setup,

numbers representing points or money are given as feedback following each selection. One possibility is that a numerical form of feedback may be easier to remember than a perceptual form of feedback such as the amount of oxygen shown in the tank. Thus, dual-task participants may be better able to remember past reward amounts if given concrete numerical reward information. For single-task participants, numerical information about the rewards given on each trial may bias them away from selecting the increasing option in the decreasing-optimal task. The decreasing option gives 90 more points on each trial than the increasing option does in this task, which should be very noticeable to participants.

However, we did not predict that changing the task setup would alter the differences between single- and dual-task participants that we found in Experiment 1. Nevertheless it is important to replicate the results in a gambling task framework to broaden the scope of our findings. Additionally, a replication would suggest that the surface features of the task do not significantly affect decision-making behavior. To our knowledge, this has not been directly tested. Such a finding might be useful for designing future work. For example, researchers might be able to examine differences in decision-making behavior in a within-subjects design by altering the surface features of the task to minimize effects of repeated testing.

Experiment 2

Method

Participants. Seventy-nine undergraduates from Texas A&M University participated in the experiment for course credit. Participants were randomly assigned to one of the four conditions that resulted from the factorial combination of two WM load (single- vs. dual-task) and two reward structure (increasing- vs. decreasing-optimal) conditions.

Materials and procedure. The reward structure for the increasing-optimal task was identical to the reward structure from the increasing-optimal task in Experiment 1, shown again in Figure 8a, and the reward structure for the decreasing-optimal task is shown in Figure 8b. These reward structures both had an end-state separation of 40 points. Figure 9 shows a sample screenshot from the experiment. Two decks were presented on each trial, a red deck and a blue deck. In the single-task condition participants selected from one of the two decks, and the resulting reward value was revealed above the chosen deck. The point total was then incremented based on the number of points earned on that trial. Participants were given a goal of earning 18,000 points in the increasing-optimal task and 22,000 points in the decreasing-optimal task. To achieve these goals, participants had to select the optimal option on about 80% of trials.

As in the first experiment, dual-task participants concurrently performed the numerical Stroop task. The procedure for the numerical Stroop task is identical to the procedure detailed in the Method section for Experiment 1 above. Participants performed a total of 250 trials and were told whether they reached their goal at the end of the experiment.

Results

Figure 10a shows the proportion of increasing option selections for participants in each condition. A 2 (WM load) \times 2 (reward

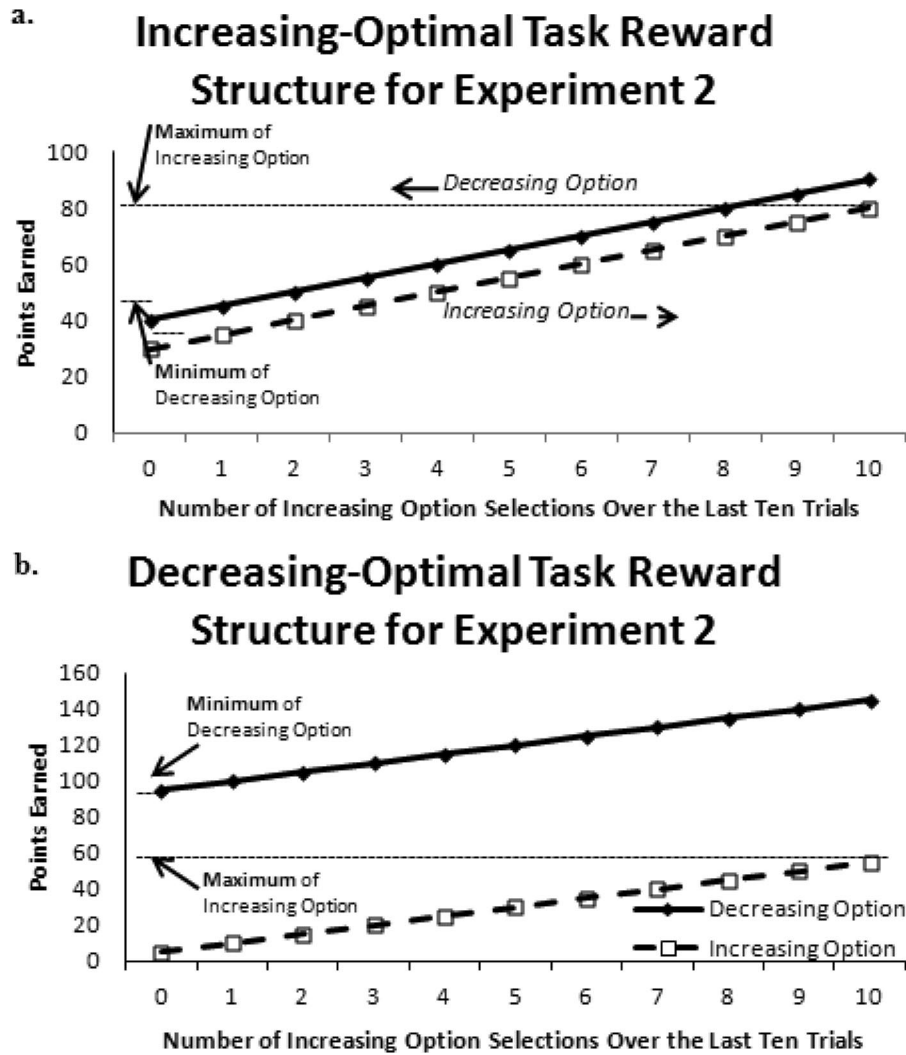


Figure 8. Rewards given for each task as a function of the number of the number of Increasing option selections over the previous 10 trials. (a) Rewards given for the increasing-optimal task. Selecting the increasing option 10 consecutive times will lead to a reward of 80 points on each trial, whereas selecting the decreasing option 10 consecutive times will lead to a reward of 40 points on each trial. (b) Rewards given for the decreasing-optimal task. Selecting the decreasing option 10 consecutive times will lead to a reward of 95 points on each trial, whereas selecting the increasing option 10 consecutive times will lead to a reward of 55 points on each trial.

structure) ANOVA revealed a significant effect of WM load, $F(1, 75) = 39.56, p < .001, \eta^2 = .35$. Single-task participants ($M = .54, SD = .27$) selected the increasing option more often than dual-task participants did ($M = .24, SD = .17$). There was also a significant effect of reward structure, $F(1, 75) = 7.18, p < .01, \eta^2 = .09$. Participants who performed the increasing-optimal task ($M = .45, SD = .27$) selected the increasing option more often than participants who performed the decreasing-optimal task did ($M = .32, SD = .25$). The WM Load \times Reward Structure interaction was not significant ($F < 1$).

Figure 10b shows the proportion of the optimal cumulative payoff earned by participants in each condition. This was computed in the same way as in Experiment 1 and is roughly equivalent to the proportion of times they selected the optimal option for

the task. A 2 (WM load) \times 2 (reward structure) ANOVA revealed a main effect of reward structure, $F(1, 75) = 25.46, p < .001, \eta^2 = .25$, where participants who performed the decreasing-optimal task ($M = .69, SD = .24$) earned a higher proportion of the optimal cumulative payoff than participants who performed the increasing-optimal task did ($M = .45, SD = .27$). The effect of WM load was nonsignificant ($F < 1$); however, there was a significant WM Load \times Reward Structure interaction, $F(1, 75) = 39.32, p < .001, \eta^2 = .34$. Pairwise comparisons within each reward structure condition showed that single-task participants ($M = .62, SD = .25$) earned a higher proportion of the optimal cumulative payoff than dual-task participants did ($M = .28, SD = .17$) in the increasing-optimal task, $F(1, 38) = 24.89, p < .001, \eta^2 = .40$, and dual-task participants ($M = .81, SD = .16$) earned

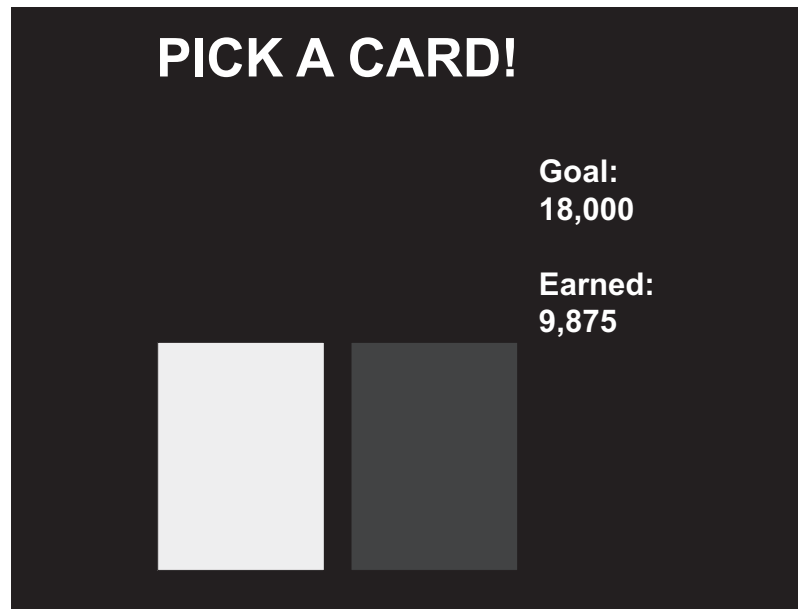


Figure 9. Sample screen shot from the gambling task used in Experiment 2. Participants received points after each selection and had a goal of earning a certain amount of points over the course of the experiment.

a higher proportion of the optimal cumulative payoff than single-task participants did ($M = .56$, $SD = .24$) in the decreasing-optimal task, $F(1, 38) = 14.94$, $p < .001$, $\eta^2 = .25$.

Figure 11 shows the average response times for participants in each condition. A 2 (WM load) \times 2 (reward structure) ANOVA revealed a main effect of WM load, $F(1, 75) = 32.50$, $p < .001$, $\eta^2 = .30$. Single-task participants ($M = .67$, $SD = .54$) took significantly less time to respond than dual-task participants did ($M = 1.32$, $SD = .46$). There was no effect of reward structure, $F(1, 75) = 1.13$, $p > .10$, and no WM Load \times Reward Structure interaction ($F < 1$).

Model-based analyses. We used the modeling approach from Experiment 1 to examine the types of strategies participants used. Table 3 lists the fits of each learning model relative to the stochastic baseline model. Positive values indicate a better fit for each learning model than the baseline model. The results are similar to those from Experiment 1. Single-task participants are better fit by the WLS model, and dual-task participants are better fit by the Softmax model. Fits of the learning models relative to the baseline model are also higher for single-task participants. In the increasing-optimal task, 10% of data sets from single-task participants are best fit by the baseline model, compared to 40% of data sets for dual-task participants ($p < .01$ by binomial test). However, in the decreasing-optimal task, the difference in the number of data sets best fit by the baseline model was much smaller, with 5% of data sets from single-task participants best fit by the baseline model, compared to 10% of data sets from dual-task participants. This was not a significant difference ($p > .10$ by binomial test). Thus, dual-task participants who performed the decreasing-optimal task did not exhibit more stochastic behavior than single-task participants who performed the same task.

Figure 12 shows the $Relative\ fit_{WLS}$ values for participants in each condition. There was a significant effect of WM load, Mann–

Whitney $U = 299.00$, $p < .001$. Single-task participants ($M = 44.62$, $SD = 58.29$) had much higher $Relative\ fit_{WLS}$ values than dual-task participants did ($M = -11.63$, $SD = 22.40$). The effect of reward structure was nonsignificant, Mann–Whitney $U = 937.00$, $p > .10$. Thus, replicating our result from Experiment 1, single-task participants were fit better by the WLS model, and dual-task participants were fit better by the Softmax model.

Parameter estimates. We next examined the best fitting parameter estimates from each model (see Table 4). We conducted a 2 (WM load) \times 2 (reward structure) ANOVA on the average estimated recency parameter (α) values for participants in each condition. There was a marginally significant effect of WM load, $F(1, 75) = 3.19$, $p < .10$, $\eta^2 = .04$. Single-task participants' data ($M = .47$, $SD = .47$) were best fit by recency parameter values higher than those for dual-task participants' data ($M = .29$, $SD = .41$). There was also a significant correlation between estimated recency parameter values and $Relative\ fit_{WLS}$ values ($r = .26$, $p < .05$). This result is consistent with the results from Experiment 1 and Otto et al. (2011), and it suggests that the Softmax model accounts for single-task participants' increased WLS strategy use by estimating higher recency parameter values for single-task participants than for dual-task participants. A 2 (WM load) \times 2 (reward structure) ANOVA on the average estimated exploitation parameter (γ) values for participants in each condition revealed no effect of WM load or reward structure ($F_s < 1$) and no WM Load \times Reward Structure interaction, $F(1, 75) = 1.89$, $p > .10$.

We next conducted a 2 (WM load) \times 2 (reward structure) ANOVA on the average estimated $P(stay|win)$ parameter values for participants in each condition. There was no effect of WM load ($F < 1$), but there was a marginally significant effect of reward structure, $F(1, 75) = 2.91$, $p < .10$, $\eta^2 = .04$. $P(stay|win)$ parameter values were higher for participants who performed the decreasing-optimal task ($M = .84$, $SD = .17$) than for participants

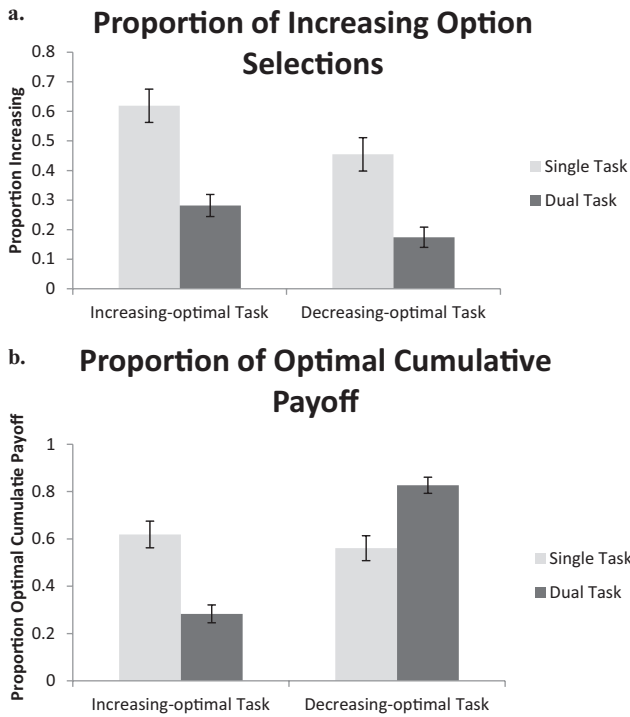


Figure 10. (a) Proportion of Increasing options selections for participants in each condition of Experiment 2. (b) Proportion of the optimal cumulative payoff participants earned in the task. These proportions are roughly equivalent to the proportion of times participants selected the optimal option based on the reward structure. Error bars represent standard errors.

who performed the increasing-optimal task ($M = .77, SD = .19$). There was also a marginally significant WM Load \times Reward Structure interaction, $F(1, 75) = 2.79, p < .10, \eta^2 = .04$. For the increasing-optimal task, single-task participants' data ($M = .81, SD = .21$) were fit best by $P(stay|win)$ parameter values higher than those for dual-task participants' data ($M = .73, SD = .17$), although this difference was nonsignificant, $F(1, 38) = 1.65, p > .10$. In contrast, for the decreasing-optimal task, single-task participants' data ($M = .81, SD = .87$) were fit best by $P(stay|win)$

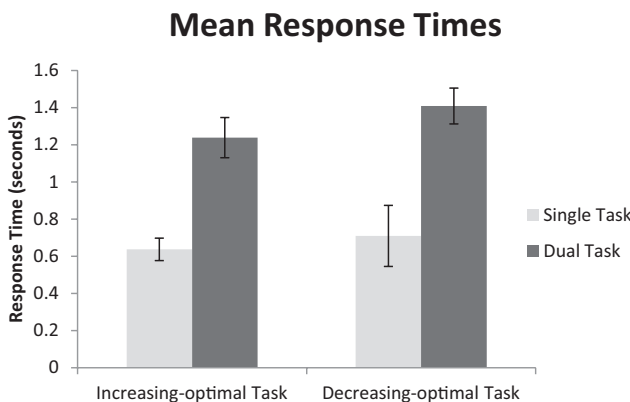


Figure 11. Average response times for participants in each condition. Error bars represent standard errors.

Table 3
Relative Fits of the WSLs and Softmax Models to the Baseline Model for Experiment 2

Task and condition	WSLS	Softmax
Increasing-optimal task		
Single-task	63.62 (12.47)	39.51 (11.83)
Dual-task	-6.13 (5.45)	3.35 (3.07)
Decreasing-optimal task		
Single-task	69.35 (12.62)	4.21 (4.76)
Dual-task	5.12 (8.09)	11.51 (3.15)

Note. Positive values indicate a better fit of each model than the baseline model. Values shown are means and standard errors of the mean (SE). WSLs = win-stay, lose-shift; Softmax = Softmax reinforcement learning.

parameter values lower than those for dual-task participants' data ($M = .87, SD = .12$), although this difference also was nonsignificant, $F(1, 37) = 1.15, p > .10$.

A 2 (WM load) \times 2 (reward structure) ANOVA on the average estimated $P(lose|shift)$ parameter values for participants in each condition showed no effect of WM load ($F < 1$), but there was a significant effect of reward structure, $F(1, 75) = 6.03, p < .05, \eta^2 = .07$. $P(lose|shift)$ parameter values were higher for participants who performed the increasing-optimal task ($M = .54, SD = .15$) than for participants who performed the decreasing-optimal task ($M = .45, SD = .16$). The WM Load \times Reward Structure interaction did not reach significance, $F(1, 75) = 1.56, p > .10$.

Cross-fitting analysis. We performed the same cross-fitting analysis that we performed for Experiment 1 data to examine the degree to which each model can mimic the other model. Figure 13 plots histograms of the $Relative\ fit_{WSLS}$ values for data generated by each model for the increasing-optimal task (Figure 13a) and the decreasing-optimal task (Figure 13b). For the increasing-optimal task, 76% of data sets generated by the Softmax model were also fit best by the Softmax model, and 98% percent of data sets generated by the WSLs model were also fit best by the WSLs model. For the decreasing-optimal task, 81% of data sets generated by the Softmax model were also fit best by the Softmax model, and 99% of data sets generated by the WSLs model were also fit best by the WSLs model. Similar to the results from Experiment 1, a

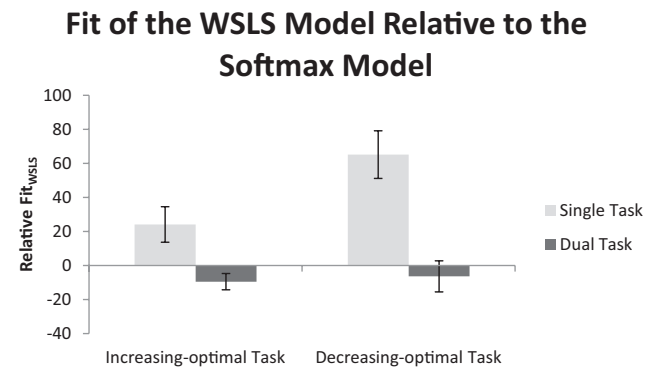


Figure 12. Average relative fit values of the WSLs model compared to the Softmax model for data from Experiment 2. Higher values indicate a better fit for the WSLs model. Error bars represent standard errors. WSLs = win-stay, lose-shift; Softmax = Softmax reinforcement learning.

Table 4
Average Best Fitting Parameter Values for Participants in Each Condition in Experiment 4

Task and condition	$P(\text{stay} \text{win})$	$P(\text{shift} \text{loss})$	α	γ
Increasing-optimal task				
Single-task	.81 (.05)	.56 (.04)	.42 (.11)	.17 (.06)
Dual-task	.73 (.04)	.52 (.02)	.29 (.07)	.07 (.01)
Decreasing-optimal task				
Single-task	.81 (.05)	.43 (.03)	.53 (.11)	.07 (.05)
Dual-task	.87 (.41)	.47 (.03)	.29 (.10)	.13 (.07)

Note. Standard errors of the mean are listed in parentheses.

large number of data sets clustered around the 0 point on the x -axis. This suggests that the WLS model provided an equal or only slightly better fit than the Softmax model to a large number of data sets that were generated by the Softmax model. Thus, the WLS model appears to be better able to mimic the Softmax model, but only to the degree that it can fit a portion of the Softmax-generated data sets roughly equally well but not substantially better.

Discussion

The pattern of results is identical to the pattern from Experiment 1. Single-task participants selected the increasing option more often than dual-task participants did, regardless of which option was optimal for the task. Dual-task participants also took a longer time to respond. This suggests that they were not simply making a quick decision to get to the numerical Stroop query but may have engaged in effortful encoding of the numerical Stroop stimuli at the beginning of each trial before making a decision. Our model comparison revealed that single-task participants were better fit by a model that assumes a rule-based WLS strategy, and dual-task participants were better fit by a Softmax reinforcement learning model that assumes a more associative strategy. When we compared the parameter estimates for participants in each condition we found that the Softmax model's recency, or learning rate, parameter values were higher for single-task participants than for dual-task participants. This effect was also found in Experiment 1 and was previously found by Otto et al. (2011) in a binary prediction task. Thus, participants who performed the decision-making task under WM load were less responsive to recent outcomes and showed almost no evidence of WLS strategy use. Instead, dual-task participants used a strategy consistent with the Softmax model, in which they tended to select the option that, on average, gave the highest immediate reward (the decreasing option). It should also be noted that a substantial proportion of dual-task participants' data were fit best by the baseline model (although this was not the case for the decreasing-optimal task in Experiment 2). This suggests that these participants may have used very little information from the rewards they received to adjust their response tendencies and instead selected the decreasing option on a large number of trials. This behavior seems plausible in that it was likely salient to the dual-task participants early in the task that the decreasing option always gave a higher immediate reward, and so they simply decided to select this option on the majority of trials. The baseline comparisons across both experiments do suggest that

dual-task participants exhibited attenuated learning in response to the rewards they received on each trial. However, we do not feel that our results suggest that dual-task participants were simply behaving randomly, as they clearly showed a preference for the decreasing option.

General Discussion

WM load is a common variable in everyday decision making. People make decisions both while focusing solely on the decision at hand and while keeping intrusive, extraneous information in mind (Burgess, 2000; Burgess, Veitch, Costello, & Shallice, 2000; Strayer & Drews, 2007). One straightforward prediction is that more available WM resources will lead to better decision making, yet our results suggest that this is not the case. They suggest that WM load affects the degree to which participants focus on the immediate versus delayed benefits of each option and whether people operate in an explicit, rule-based mode that utilizes heuristics (like WLS) or whether they operate in a more implicit, associative mode that maximizes the options with the highest immediate reward. Participants who performed the task under single-task conditions did not simply make better decisions but instead made their decisions in a different manner than participants who were placed under WM load did. Single-task participants made decisions based on a comparison of the current reward to the reward received on the previous trial, and dual-task participants made decisions based on which option would give the best immediate reward. Focusing on the differences in rewards received across trials was productive for one task (the increasing-optimal task) but counterproductive for the other task (the decreasing-optimal task).

When placed in the context of the three components of decision making outlined above, our results suggest that WM load does not affect the third, and perhaps most important, component of decision making, the juxtaposition of the immediate versus delayed consequences of selecting each option. Rather, WM load appears to directly affect the first two components. The results for participants in the single-task conditions of each experiment are perhaps most interesting. Single-task participants exhibited sensitivity to the increasing option's effect on the rewards for both options on future trials, yet they seemed to pick this option without realizing that the immediate benefits of the decreasing option outweighed the delayed benefits of the increasing option in the decreasing-optimal tasks. This pattern of results generalized across two separate experiments with different reward structures (small vs. large end-state separation) and different surface features (Mars farming and gambling). Single-task participants selected the increasing option more often than dual-task participants did, even when the decreasing option gave 90 more points than the increasing option on each trial and when the difference in end-state separation was 40 points (Experiment 2).

Moreover, the decreasing-optimal task can be considered "easier" than the increasing-optimal task because participants earned a higher proportion of the optimal cumulative payoff in the decreasing-optimal conditions in both experiments. Thus, single-task participants performed better on the "harder" increasing-optimal tasks but worse on the "easier" decreasing-optimal tasks. These results could suggest a lack of metacognitive awareness among our participants, in that they were unable to appropriately

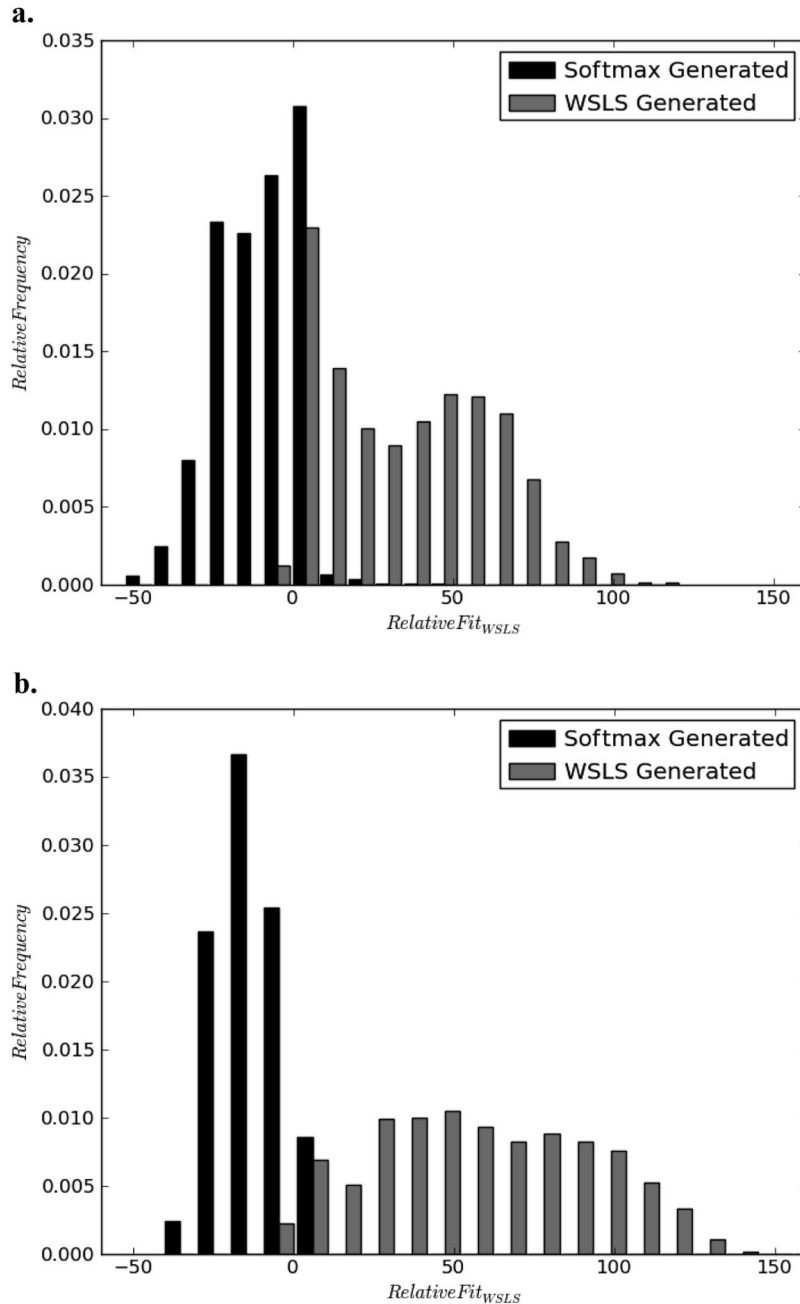


Figure 13. Distributions of fits of the WSL model relative to the Softmax model for data simulated by each model for the (a) increasing-optimal and (b) decreasing-optimal tasks. Values greater than zero indicate a better fit for the WSL mode, values less than zero indicate a better fit for the Softmax model, and values equal to zero indicate an equal fit for the two models. WSL = win-stay, lose-shift; Softmax = Softmax reinforcement learning.

weigh the immediate and delayed benefits of each option to make the best decisions.

One caveat to this work is that we chose one working memory load manipulation (the numerical Stroop task) among many. Other working memory tasks, such as memorizing a digit span, keeping a running count of tones presented during the task, or generating a random number, have also been successfully used to manipulate

WM load. Although we believe that the numerical Stroop task successfully depleted working memory resources, there is also evidence that different WM load manipulations can differentially affect working memory capacity (e.g., Jameson, Hinson, & Whitney, 2004; Miles & Minda, 2011). Thus, the possible differences in the effects of specific WM load manipulations should be considered when designing future research.

Dual Learning Systems

The implications of these results should also be considered for dual-system theories cognition. Many dual-system theories would likely predict that worse performance for dual-task participants on the increasing-optimal task was due to those participants using a more reflexive strategy, which would favor options with larger immediate rewards. However, one could predict that single-task participants who are operating in a more reflective, or “cold-cognitive,” mode should be capable of more rational, abstract reasoning and should be able to learn the optimal choice for the task (Evans, 2003).

Our modeling results suggest that, rather than behaving rationally, single-task participants used their unoccupied WM resources to employ a simple, rule-based strategy that was not always optimal. Future work should examine what affects this third opponent of decision making. One possibility is that, for single-task participants, the two systems are operating simultaneously and that the explicit system is dominant. This view of competition between the systems is a central feature of Ashby and colleagues’ Competition between Verbal and Implicit Systems (COVIS) theory of category learning (Ashby et al., 1998). One possibility is that the implicit system interferes with the explicit system under single-task conditions, which leads single-task participants to utilize simple rules such as WLS rather than reflectively juxtapose the strengths and weaknesses of each option. We admit that this line of reasoning is speculative, but future work could potentially develop testable predictions from theories that propose competition between implicit and explicit systems.

As we acknowledged in the introduction, the notion that the Softmax model mimics an implicit or procedural processing mode is consistent with work that has demonstrated that subcortical regions can develop and update expected reward values based on the immediate rewards received after selecting each option (Hare et al., 2008; Pagnoni et al., 2002; Pessiglione et al., 2006). However, the strategy “select the option that has, on average, given the largest rewards on recent trials” is obviously verbalizable and capable of being explicitly represented. This strategy is likely more simple than the WLS shift strategy, and so the dual task may have simply led participants to use a simpler strategy that is also explicit in nature.

Related to the dual-systems issue is the role that affect played in guiding our participants’ behavior. Many dual-systems theories propose that the implicit system is the “hot” or affectively rich system (Evans, 2003; Sloman, 1996), and affect has been shown to play a role in gambling tasks like the ones our participants performed (Bechara, Damasio, & Damasio, 2000; Bechara et al., 1994). One possibility is that the hot, affective system favors options based on their immediate rewards and that the cold-cognitive system must override this tendency to select options that do not lead to the best immediate rewards. This notion of competition between implicit (affective) and explicit (cognitive) systems, whereby the explicit system must override the implicit system, is a central feature of many dual-system theories (e.g., Cohen, 2005; Sloman, 1996). In our experiments, the compromised working memory resources of dual-task participants may have prevented them from overriding the affect-based tendency to select the option that gave higher immediate rewards.

Implications and Future Directions

A central aim for future work should be to examine what processes affect the third component of decision making outlined above, which could be described as making the optimal decision for the situation. One possibility is that WM capacity affects the juxtaposition of the immediate and long-term consequences of each action, but WM load does not. Our study examined only WM load and did not measure individual differences in WM capacity. Future work should identify whether individual differences in WM capacity affect decision making in situations similar to the experiments reported here and also whether WM capacity interacts with WM load. Individuals with high WM capacity may outperform individuals with low WM capacity on both tasks; however, it is also possible that individual differences in WM capacity will have the same effect on preferences for immediate versus delayed outcomes that we found by manipulating WM load.

These results have important implications for everyday decision making. Multitasking has become a ubiquitous part of modern life (e.g., Burgess et al., 2000; Vestergren & Nilsson, 2011). Here we demonstrated that the degree to which WM resources are divided between separate, simultaneous tasks appears to have very large effects on the strategies people employ in decision making. Our results indicate that attenuating WM resources by performing a concurrent task can lead to a temporal myopia in decision making whereby individuals are able to focus on only the immediate consequences of their actions. Intriguingly, our results also suggest that being free from a distracting secondary task is not always beneficial. The full resources available to our single-task participants may have led them to employ a more WM-demanding strategy than was needed for a relatively easy task.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. doi:10.1109/TAC.1974.1100705
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442–481. doi:10.1037/0033-295X.105.3.442
- Ashby, F. G., & Ell, S. W. (2002). Single versus multiple systems of category learning: Reply to Nosofsky & Kruschke (2002). *Psychonomic Bulletin & Review*, *9*, 175–180. doi:10.3758/BF03196275
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149–178. doi:10.1146/annurev.psych.56.091103.070217
- Bechara, A., Damasio, A. R., & Damasio, H. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*, 295–307. doi:10.1093/cercor/10.3.295
- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15. doi:10.1016/0010-0277(94)90018-3
- Bechara, A., & Martin, E. M. (2004). Impaired decision making related to working memory deficits in individuals with substance addictions. *Neuropsychology*, *18*, 152–162. doi:10.1037/0894-4105.18.1.152
- Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research*, *63*, 279–288. doi:10.1007/s004269900006
- Burgess, P. W., Veitch, E., Costello, A. D. L., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*, 848–863. doi:10.1016/S0028-3932(99)00134-7

- Cohen, J. D. (2005). The vulcanization of the human brain: A neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives*, *19*, 3–24. doi:10.1257/089533005775196750
- Curtis, C. E., & Lee, D. (2010). Beyond working memory: The role of persistent activity in decision making. *Trends in Cognitive Sciences*, *14*, 216–222. doi:10.1016/j.tics.2010.03.006
- Daw, N. D., O'Doherty, J. P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*, 876–879.
- Djuric, P. M. (1998). Asymptotic MAP criteria for model selection. *IEEE Transactions on Signal Processing*, *46*, 2726–2735. doi:10.1109/78.720374
- Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E. J. (2011). Diffusion versus linear ballistic accumulation: Different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, *18*, 61–69. doi:10.3758/s13423-010-0022-4
- Dretsche, M. N., & Tipples, J. (2008). Working memory involved in predicting future outcomes based on past experiences. *Brain and Cognition*, *66*, 83–90. doi:10.1016/j.bandc.2007.05.006
- Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences*, *7*, 454–459. doi:10.1016/j.tics.2003.08.012
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. doi:10.1146/annurev.psych.59.103006.093629
- Filoteo, J. V., Lauritzen, S., & Maddox, W. T. (2010). Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychological Science*, *21*, 415–423. doi:10.1177/0956797610362646
- Franco-Watkins, A. M., Pashler, H., & Rickard, T. C. (2006). Does working memory load lead to greater impulsivity? Commentary on Hinson, Jameson, & Whitney's (2003). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 443–447. doi:10.1037/0278-7393.32.2.443
- Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review*, *113*, 300–326. doi:10.1037/0033-295X.113.2.300
- Frank, M. J., & Kong, L. (2008). Learning to avoid in older age. *Psychology and Aging*, *23*, 392–398. doi:10.1037/0882-7974.23.2.392
- Frank, M. J., O'Reilly, R. C., & Curran, T. (2006). When memory fails, intuition reigns: Midazolam enhances implicit inference in humans. *Psychological Science*, *17*, 700–707. doi:10.1111/j.1467-9280.2006.01769.x
- Gureckis, T. M., & Love, B. C. (2009a). Learning in noise: Dynamic decision-making in a variable environment. *Journal of Mathematical Psychology*, *53*, 180–193. doi:10.1016/j.jmp.2009.02.004
- Gureckis, T. M., & Love, B. C. (2009b). Short-term gains, long term pains: How cues about state aid in learning in dynamic environments. *Cognition*, *113*, 293–313. doi:10.1016/j.cognition.2009.03.013
- Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience*, *28*, 5623–5630. doi:10.1523/JNEUROSCI.1309-08.2008
- Hinson, J. M., Jameson, T. L., & Whitney, P. (2002). Somatic markers, working memory, and decision making. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 341–353. doi:10.3758/CABN.2.4.341
- Jameson, T. L., Hinson, J. M., & Whitney, P. (2004). Components of working memory and somatic markers in decision making. *Psychonomic Bulletin & Review*, *11*, 515–520. doi:10.3758/BF03196604
- Lee, M. D., Zhang, S., Munro, M., & Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research*, *12*, 164–174. doi:10.1016/j.cogsys.2010.07.007
- Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning systems of perceptual category learning. *Behavioral Processes*, *66*, 309–332. doi:10.1016/j.beproc.2004.03.011
- Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & Cognition*, *32*, 582–591. doi:10.3758/BF03195849
- Metcalfe, J., & Mischel, W. (1999). A hot/cool system analysis of delay of gratification: Dynamics of willpower. *Psychological Review*, *106*, 3–19. doi:10.1037/0033-295X.106.1.3
- Miles, S. J., & Minda, J. P. (2011). The effect of concurrent verbal and visual tasks on category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 588–607. doi:10.1037/a0022309
- Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, *4*, 79–95. doi:10.3758/BF03210778
- Newell, B. R., Dunn, J. C., & Kalish, M. (2010). The dimensionality of perceptual category learning: A state-trace analysis. *Memory & Cognition*, *38*, 563–581. doi:10.3758/MC.38.5.563
- Nosofsky, R. M., & Kruschke, J. K. (2002). Single-system models and interference in category learning: Commentary on Waldron and Ashby (2001). *Psychonomic Bulletin & Review*, *9*, 169–174. doi:10.3758/BF03196274
- Nowak, M., & Sigmund, K. (1993, July 1). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, *364*, 56–58. doi:10.1038/364056a0
- Otto, A. R., Gureckis, T. M., Markman, A. B., & Love, B. C. (2009). Navigating through abstract decision spaces: Evaluating the role of state generalization in a dynamic decision-making task. *Psychonomic Bulletin & Review*, *16*, 957–963. doi:10.3758/PBR.16.5.957
- Otto, A. R., Markman, A. B., Gureckis, T. M., & Love, B. C. (2010). Regulatory fit and systematic exploration in a dynamic decision-making environment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 797–804. doi:10.1037/a0018999
- Otto, A. R., Markman, A. B., & Love, B. C. (2012). Taking more, now: The optimality of impulsive choice hinges on environment structure. *Social Psychological & Personality Science*, *3*(2), 131–138. doi:10.1177/1948550611411311
- Otto, A. R., Taylor, E. G., & Markman, A. B. (2011). There are at least two kinds of probability matching: Evidence from a secondary task. *Cognition*, *118*, 274–279. doi:10.1016/j.cognition.2010.11.009
- Pagnoni, G., Zink, C. F., Montague, P. R., & Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*, *5*, 97–98. doi:10.1038/nn802
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006, August 31). Dopamine-dependent prediction errors underpin reward-seeking behavior in humans. *Nature*, *442*, 1042–1045. doi:10.1038/nature05051
- Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*, 245–251. doi:10.1016/S0028-3932(02)00157-4
- Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14. doi:10.1002/bdm.681
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22. doi:10.1037/0033-2909.119.1.3
- Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131. doi:10.1207/S15327957PSPR0402_01
- Steyvers, M., Lee, M. D., & Wagenmakers, E. J. (2009). A Bayesian

- analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, 53, 168–179. doi:10.1016/j.jmp.2008.11.002
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247. doi:10.1207/s15327957pspr0803_1
- Strayer, D. L., & Drews, F. A. (2007). Cell-phone-induced driver distraction. *Current Directions in Psychological Science*, 16, 128–131. doi:10.1111/j.1467-8721.2007.00489.x
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. doi:10.1037/h0054651
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Tversky, A., & Kahneman, D. (1981, January 30). The framing of decisions and the psychology of choice. *Science*, 211, 453–458. doi:10.1126/science.7455683
- Vestergren, P., & Nilsson, L. G. (2011). Perceived causes of everyday memory problems in a population-based sample aged 39–99. *Applied Cognitive Psychology*, 25, 641–646. doi:10.1002/acp.1734
- Wagenmakers, E. J., Ratcliff, R., Gomez, P., & Iverson, G. (2004). Assessing model mimicry using the parametric bootstrap. *Journal of Mathematical Psychology*, 48, 28–50. doi:10.1016/j.jmp.2003.11.004
- Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, 8, 168–176. doi:10.3758/BF03196154
- Worthy, D. A., Gorlick, M. A., Pacheco, J. L., Schnyer, D. M., & Maddox, W. T. (2011). With age comes wisdom: Decision-making in younger and older adults. *Psychological Science*, 22, 1375–1380. doi:10.1177/0956797611420301
- Worthy, D. A., & Maddox, W. T. (2012). Age-based differences in strategy use in choice tasks. *Frontiers in Neuroscience*, 5, 145–155.
- Worthy, D. A., Maddox, W. T., & Markman, A. B. (2007). Regulatory fit effects in a choice task. *Psychonomic Bulletin & Review*, 14, 1125–1132. doi:10.3758/BF03193101
- Worthy, D. A., Maddox, W. T., & Markman, A. B. (2008). Ratio and difference comparisons of expected reward in decision-making tasks. *Memory & Cognition*, 36, 1460–1469. doi:10.3758/MC.36.8.1460
- Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review*, 12, 387–402. doi:10.3758/BF03193783
- Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science*, 16, 973–978. doi:10.1111/j.1467-9280.2005.01646.x
- Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, 34, 387–398. doi:10.3758/BF03193416

Received July 14, 2011

Revision received February 26, 2012

Accepted March 5, 2012 ■