



Published in final edited form as:

Cogn Affect Behav Neurosci. 2014 December ; 14(4): 1208–1220. doi:10.3758/s13415-014-0293-8.

State-Based Versus Reward-Based Motivation in Younger and Older Adults

Darrell A. Worthy¹, Jessica A. Cooper², Kaileigh A. Byrne¹, Marissa A. Gorlick², and W. Todd Maddox²

¹Texas A&M University

²University of Texas at Austin

Abstract

Recent decision-making work has focused on a distinction between a habitual, *model-free* neural system that is motivated toward actions that lead directly to reward, and a more computationally demanding goal-directed, *model-based* system that is motivated toward actions that improve one's future state. In this paper we examine how aging affects motivation toward reward-based versus state-based decision-making. Participants performed tasks in which one type of option provided larger immediate rewards, but the alternative type of option led to larger rewards on future trials, or improvements in state. We predicted that older adults would show a reduced preference for choices that led to improvements in state and a greater preference for choices that maximized immediate reward. We also predicted that fits from a HYBRID reinforcement-learning model would indicate greater model-based strategy use in younger than in older adults. In line with these predictions, older adults selected the options that maximized reward more often than younger adults in three of the four tasks and modeling results suggested reduced model-based strategy-use. In the task where older adults showed similar behavior to younger adults our model-fitting results suggested that this was due to the utilization of a win-stay-lose-shift heuristic rather than a more complex model-based strategy. Additionally, within older adults we found that model-based strategy use was positively correlated with memory measures from our neuropsychological test battery. We suggest that this shift from state-based to reward-based motivation may be due to age related declines in the neural structures needed for more computationally demanding model-based decision-making.

Individuals of all ages must regularly make important decisions and a central component of decision-making is motivation. Individuals are motivated to make decisions that are more likely to lead to positive outcomes and less likely to lead to negative ones. Recently, there has been a surge of work aimed at examining the distinction between model-free versus model-based decision-making strategies (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Gershman, Markman, & Otto, 2012; Gläscher, Daw, Dayan, & O'Doherty, 2010; Otto, Gershman, Markman, & Daw, 2013; Otto, Raio, Chiang, Phelps, & Daw, 2013). Motivation plays a prominent role in distinguishing these two approaches to decision-making situations. Model-free decision-making is habitual and the motivational focus is centered on

performing actions that lead to reward or avoid punishment. Actions that lead to reward are reinforced and actions that lead to either punishment or no reward are not. By contrast, model-based decision-making is goal-directed and the motivational focus is on performing actions that improve one's future state. Model-based decision-making requires individuals to utilize a model of the environment and consider how each action can affect both immediate and future outcomes. Model-free decision-making is *reward-based* because individuals are primarily motivated to perform actions that are followed by reward, while model-based decision-making is *state-based* because individuals are primarily motivated to perform actions that improve their future state (Gläscher et al., 2010).

Model-based and model-free decision-making processes are thought to be mediated by separate neural systems, with the weight given to each system varying across individuals and under different circumstances. Areas of the ventral striatum and medial prefrontal cortex (mPFC) are thought to be critical for model-free decision-making (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; O'Doherty, 2004). In addition to ventral striatal regions, the intraparietal sulcus and lateral regions of the PFC, particularly the dorsolateral PFC (DLPFC) have been implicated in model-based decision-making (Gläscher et al., 2010; Smittenaar, FitzGerald, Romei, Wright, & Dolan, 2013). For example, Smittenaar and colleagues showed that impairment of right DLPFC via transcranial magnetic stimulation (TMS) reduced model-based responding (Smittenaar et al., 2013). However, one study found that the model-based and model-free systems may utilize some of the same neural structures to compute state and reward-based information, particularly the ventral striatum (Daw et al., 2011).

In addition to neural dissociations of model-based and model-free systems, another distinction is that model-based decision-making is more computationally demanding as it requires a representation of the environment as well as internal goal states. Several recent studies have found an association between model-based decision-making and working memory processes (Gershman et al., 2012; Otto, Gershman, et al., 2013; Otto, Raio, et al., 2013). In one study Otto and colleagues found that a stress manipulation shifted the balance from model-based to model-free decision-making and that individuals who were lower in working memory capacity were more likely to utilize model-free strategies when placed under stress (Otto, Raio, et al., 2013).

One important issue to consider is how aging affects this balance between model-based and model-free decision-making. Another way of framing this issue is whether healthy aging affects the degree to which individuals are motivated by habitual, reward-based processes versus goal-directed, state-based processes. One possibility is that older adults will show reduced model-based learning compared to younger adults due to it being more computationally demanding. There is extensive evidence that aging is associated with declines in attention, working memory, and executive control (Salthouse, 2004, 2009). Given the increased cognitive demand for model-based compared to model-free decision-making a clear prediction is that older adults will be less likely to engage in model-based decision-making. Eppinger and colleagues recently found support for this hypothesis (Eppinger, Walter, Heekeren, & Li, 2013). They utilized a two-stage Markov decision-making task and found that younger adults were more likely to use a model-based strategy

than older adult who showed more evidence of model-free strategy-use. Compared to younger adults, older adults were more influenced by rewards and showed less strategic exploration of the task structure. The reduced reliance on model-based responding also led to poorer performance in terms of overall payoffs accumulated in the task. These results are consistent with other work that has found decision-making deficits in older adults (Denburg, Tranel, & Bechara, 2005; Denburg et al., 2009; Samanez-Larkin, Kuhnen, Yoo, & Knutson, 2010).

However, in a recent paper from our lab we obtained results that we then interpreted as evidence for *enhanced* model-based responding in older adults (Worthy, Gorlick, Pacheco, Schnyer, & Maddox, 2011). In this study, older adults were more likely than younger adults to forego an option that led to larger immediate payoffs on each trial, and instead select an option that led to smaller immediate payoffs on each trial but also led to larger rewards on future trials, which made it the best choice. We concluded that this result was due to older adults engaging in a more model-based decision-making strategy where the future outcomes of their actions were given greater weight than the immediate outcomes.

One potential way to reconcile our findings with the recent results from Eppinger and colleagues (Eppinger, Walter, et al., 2013), that showed reduced model-based decision-making in older adults, comes from another study in which we found that older adults are more likely to employ a heuristic-based win-stay-lose-shift (WSLS) strategy than younger adults (Worthy & Maddox, 2012). WSLS is a simple heuristic-based strategy where participants show a certain propensity to “stayby picking the same option following improvements in reward, or “shift” to a different option following declines in reward (Otto, Taylor, & Markman, 2011; Worthy, Hawthorne, & Otto, 2013; Worthy, Otto, & Maddox, 2012). A WSLS strategy is distinct from both model-based and model-free strategies which are both reinforcement learning (RL) strategies that assume that participants probabilistically select options with higher expected values (Otto et al., 2011; Sutton & Barto, 1998; Worthy et al., 2013; Worthy et al., 2012). Older adults may show behavior that suggests that they are utilizing a complex model-based strategy, when in fact they are employing a simple heuristic-based WSLS strategy that is less computationally demanding.

In the current work we examine the degree to which older and younger adults utilize model-free versus model-based reinforcement learning strategies as well as a heuristic-based WSLS strategy. We examine this issue in different tasks for which the optimal strategy is to either forego more immediately rewarding options in favor of options that provide larger rewards on future trials (Increasing-optimal task), or to select the larger immediately rewarding options because the gain from selecting the other options which increase rewards on future trials is so small that selecting the options that provide larger immediate rewards is the optimal strategy (Decreasing-optimal task). For each task, participants perform either a two or four-choice variant. The two-choice tasks in each of these experiments are direct replications of our prior work where older adults outperformed younger adults (Worthy et al., 2011). The four-choice variants extend these findings as they explore more computationally demanding decisions. Additionally, while prior work suggests that a WSLS strategy can account for performance in the two choice tasks (Worthy et al., 2012), as we

detail below, it will likely be less efficient when deciding from amongst four alternatives compared to model-free or model-based reinforcement-learning strategies.

Recent work has demonstrated that older adults actually prefer fewer decision options in a variety of contexts (Reed, Mikels, & Simon, 2008). While more options may seem like a beneficial or preferable aspect of many decision-making situations it can also unnecessarily increase the difficulty in obtaining and processing information about the various alternatives (Schwartz, 2000, 2009). While older adults have shown reduced preference for more decision options (Reed et al., 2008), and the number of choice options is a fairly pervasive variable in many important decision-making situations, to our knowledge, no studies have directly examined how the number of decision options affects decision-making behavior in older and younger adults.

In the remainder of the paper we first present our experiment and the behavioral results. We then fit the data with a simple WSLs model and a sophisticated HYBRID RL model that allows us to infer the degree to which older and younger adults are relying on model-free versus model-based processes. We predict that the RL model will fit the data best overall, particularly in the four-choice conditions, and that older adults will show less evidence of model-based responding than younger adults. However, we also predict that older adults will show more evidence of utilizing a WSLs strategy than younger adult in the two-choice Increasing optimal task which should lead to equal or better performance in the task, despite differences in strategy use.

Experiment

Participants completed a dynamic decision-making task that differed based on the combination of two task types (Increasing or Decreasing optimal) and two different numbers of choice alternatives (two-choices versus four-choices). These types of dynamic decision-making tasks have been used in much recent work to examine how people learn to forego larger immediate rewards in favor of options that improve future rewards (Byrne & Worthy, 2013; Cooper, Worthy, Gorlick, & Maddox, 2013; Gureckis & Love, 2009a, 2009b; Maddox, Gorlick, Worthy, & Beevers, 2012; Otto, Gureckis, Markman, & Love, 2009; Worthy et al., 2011). Figure 1 shows the reward structure for each task. In the two-choice variant there are two options: a decreasing option and an increasing option. The decreasing option always provides a larger reward on any given trial; however, the rewards possible for both options increase as the increasing option is chosen more frequently over a span of 10 trials. As participants select the increasing option more often, rewards increase, whereas as participants selected the decreasing option more often, rewards decreased. Thus, selecting the decreasing option results in greater immediate reward, but selecting the increasing option improves one's future state by leading to larger rewards for both options. The four-choice tasks were identical to the two-choice tasks except there were two increasing and two decreasing options, rather than just one of each type.

In the "Increasing-optimal" task shown in Figure 1a the optimal strategy was to consistently select the increasing option which improved participants' state by increasing rewards on future trials for both options as it was selected more often. If participants selected the

increasing option on each trial then they would eventually reach the highest state (10) and earn 80 units of oxygen on each trial throughout the task, however, if they selected the decreasing option on each trial then they would reach the lowest state (0) and earn only 40 units of oxygen on each trial. A model-based strategy should lead to better performance in the Increasing-optimal tasks compared to a model-free strategy because participants using a model-based strategy should be more likely to select the increasing option which improves their state on future trials. However, participants may select the Increasing option more frequently if they are utilizing a WSLs strategy, particularly in the two choice variant, where declines in successive rewards from selecting the Decreasing option may make them more likely to shift to the Increasing option.

In contrast, in the “Decreasing-optimal” task in Figure 1b the optimal strategy is to repeatedly select the decreasing option because the maximum value that could be obtained from repeatedly selecting the increasing option (55 units of oxygen) is smaller than the minimum value that could be obtained from simply selecting the decreasing optimal task on each trial (65 units of oxygen). A model-based strategy may actually lead to poorer performance in the Decreasing optimal task because attempting to improve one’s future state is counterproductive, and the optimal strategy is to instead select the option that leads to the largest reward on each trial.

If older adults are more likely to utilize a model-free strategy, compared to younger adults, then they should perform better on the Decreasing optimal tasks, but worse on the Increasing optimal tasks. However, our previous work suggests that older adults are able to perform as well as younger adults in the two-choice variant of the Increasing-optimal task and this may be due to their utilization of a WSLs strategy (Worthy et al., 2011; Worthy & Maddox, 2012). Fits of our HYBRID RL and WSLs models should help distinguish between model-based and model-free RL strategies and WSLs heuristic-based strategies.

Method

Participants

Ninety-one older adults (average age 67.63) from the greater Austin and College Station, Texas communities and 91 younger adults (average age 22.51) from the University of Texas and Texas A&M University communities were paid \$10 per hour for their participation. Informed consent was obtained from all participants and the experiment was approved for ethics procedures using human participants.

Neuropsychological Testing Procedures

Older adults were given a series of standardized neuropsychological tests designed to assess general intellectual ability across *attention* (WAIS-III Digit Span, (Wechsler, 1997)), *executive functioning* (Trail Making Test A&B (TMT), FAS (Lezak, 1995); Wisconsin Card Sorting Task (WCST)(Heaton, 1981), and *memory* (California Verbal Learning Test (CVLT, (Fridlund & Delis, 1987)). The tests were administered in one two-hour session.

Normative scores for each subject were calculated for each neuropsychological test using the standard age-appropriate published norms. Tables 1 and 2 show the means, standard

deviations, and ranges of standardized z-scores on each test for older adults in both the Increasing and Decreasing-optimal tasks. All WAIS subtest percentiles were calculated according to the testing instructions and then converted to standardized z-scores. The CVLT and WCST standardized t-scores were calculated according to testing directions then converted to standardized z-scores, and the TMT standard z-scores were calculated according to the testing instructions. Subjects were excluded from participation if they scored more than two standard deviations below the standardized mean on more than one neuropsychological test in the same area (memory, executive functioning, or attention). We did not have to exclude any subjects based on this criterion.

Materials and Procedure

As described above, participants completed either the two or four-choice variant of either the Increasing or Decreasing optimal task. Figure 2 shows a sample screen shot for the two-choice (Figure 2a) and four-choice (Figure 2b) conditions. On each trial, participants selected from one of the options and the narrow bar labeled “Current” would fill up to represent the amount of oxygen that had just been extracted (colored blue). The oxygen would then appear to be transferred to the larger tank labeled “Cumulative” which would indicate the amount of oxygen the participant had gained up to that point. Participants were given a goal of extracting a certain amount of oxygen by the end of the experiment, represented by the line near the top of the Cumulative tank. The goal could be achieved by selecting the optimal option on 80% of trials.

Figure 1 shows the reward structure for each task. In the two-choice task there were two options: a decreasing option and an increasing option. In the Increasing optimal task the optimal strategy was to select the increasing option on each trial, and in the Decreasing optimal task the optimal strategy was to select the Decreasing option on each trial. The four-choice task variants were identical to the two-choice variants except there were two increasing and two decreasing options, rather than just one of each type. The increasing and decreasing options were yoked so that the participant’s location on the x-axis in Figures 1a and 1b, or their state, on each trial was determined from the number of times they had selected *either* increasing option over the ten previous trials. Participants were told nothing about the reward structure of the task and had to learn the immediate and delayed effects of selecting each option from experience. They performed a total of 250 trials and were told whether or not they reached their goal at the end of the experiment.

Results

We first examined the proportion of trials where participants selected the increasing option in 50-trial blocks of the task. Selecting this option should increase participants’ state on future trials and is the optimal choice for the Increasing optimal task but the sub-optimal choice for the Decreasing optimal task. Figure 3 shows the proportion of increasing option selections separately for each task. A 2 (Age: older versus younger adults) X 2 (Task-type: Increasing versus Decreasing optimal) X 2 (Choice options: two versus four) X 5 (50-trial block) repeated measures ANOVA revealed a significant effect of block, $F(4, 696)=26.51$, $p<.001$, partial $\eta^2=.13$. There was also a significant Age X Block interaction, $F(4,696)=2.59$,

$p < .05$, partial $\eta^2 = .02$. To examine this interaction we looked at the effect of block within each age group. There was a significant effect of block for both older adults, $F(4,360) = 6.96$, $p < .001$, partial $\eta^2 = .07$, and younger adults, $F(4,360) = 20.91$, $p < .001$, partial $\eta^2 = .19$. Notably, the effect size was substantially larger for younger than for older adults. Visual inspection of Figure 3 suggests that younger adults selected the increasing option more frequently as the task progressed than older adults. There was also a significant task type X block interaction, $F(4, 696) = 7.57$, $p < .001$, partial $\eta^2 = .05$. To determine the locus of this interaction we examined the effect of block within each task type. The effect of block was significant in both the Increasing, $F(4,340) = 25.95$, $p < .001$, partial $\eta^2 = .23$, and Decreasing optimal tasks, $F(4,380) = 4.94$, $p < .01$, partial $\eta^2 = .05$, but the effect was substantially larger in the Increasing optimal task. No other interactions reached significance.

We also observed a significant between-subjects main effect of age, $F(1,174) = 13.42$, $p < .001$, partial $\eta^2 = .07$, where older adults ($M = .35$, $SD = .24$) selected the Increasing option less than younger adults overall ($M = .47$, $SD = .26$), and a significant main effect of task type, $F(1,174) = 43.48$, $p < .001$, partial $\eta^2 = .20$, where participants selected the increasing option more often in the Increasing optimal task (Increasing optimal, $M = .52$, $SD = .26$; Decreasing optimal, $M = .32$, $SD = .20$). In addition to a significant two-way interaction between age and choice options, $F(1,174) = 5.13$, $p < .05$, partial $\eta^2 = .03$, there was also a significant three-way age X task type X choice options interaction, $F(1,174) = 4.74$, $p < .05$, partial $\eta^2 = .03$.

To decompose the three-way interaction we examined effects of age and choice options within each task type condition. In the Increasing optimal task the effects of age, $F(1,82) = 2.74$, $p = .10$, and choice options did not reach significance, $F(1,82) = 1.52$, $p = .22$, but there was a significant age X choice option interaction, $F(1,82) = 7.32$, $p < .01$, partial $\eta^2 = .08$. Within the two choice condition older adults ($M = .52$, $SD = .21$) selected the Increasing option more than younger adults ($M = .45$, $SD = .30$) over all trials, but the difference was not significant, $F(1,37) < 1$, $p = .39$. In contrast, within the four-choice condition there was a significant main effect of age, $F(1,45) = 10.92$, $p < .01$, partial $\eta^2 = .20$, where younger adults ($M = .68$, $SD = .22$) selected the increasing option more frequently than older adults ($M = .44$, $SD = .26$).

In the Decreasing optimal task there was a strong effect of age, $F(1,92) = 14.97$, $p < .001$, partial $\eta^2 = .14$, where older adults ($M = .23$, $SD = .16$) selected the increasing option much less often than younger adults ($M = .39$, $SD = .21$). The effect of choice options, $F(1,92) = 1.73$, $p = .19$ and the age X choice option interaction were both non-significant ($F < 1$, $p = .94$). Thus older adults performed better on both versions of the task than younger adults.

Model-based Analyses

We fit a WSLS model as well as a HYBRID RL model similar to that recently used by Eppinger and colleagues to examine the degree to which decisions are model-free versus model-based (Eppinger, Walter, et al., 2013). In addition we also fit a Baseline or null model that assumes a stochastic response process (Worthy & Maddox, 2012; Worthy et al., 2012; Yechiam & Busemeyer, 2005).

The WSLS model is identical to the model we used in a previous paper (Worthy & Maddox, 2012) The model has two free parameters. The first parameter represents 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:

$$P(a_i, t | choice_{t-1} = a_i \& r(t-1) \geq r(t-2)) = P(stay | win). \quad (1)$$

In Equation 1 r represents the reward received on a given trial. The probability of switching to another option following a win trial is $1 - P(stay | win)$. In the four choice tasks, to determine the probability of selecting each of the other three options this value is divided by three so that the probabilities for selecting each option sum to one. This is important as it makes the WSLS strategy less effective in the four-choice task than in the two choice task because the model assumes that one of the other three alternatives is selected randomly.

The second parameter represents 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:

$$P(a_j, t | choice_{t-1} = a_i \& r(t-1) < r(t-2)) = P(shift | loss). \quad (2)$$

In the four choice task variants this probability is divided by three and assigned to each of the other three options. The probability of staying with an option following a “loss” is $1 - P(shift | loss)$. Thus, this model assumes a simple, heuristic-based strategy that requires only the reward received on the previous trial to be maintained in working memory e.g. (Otto et al., 2011; Worthy et al., 2012).

The HYBRID RL model assumes that participants observed the hidden state (s) on each trial, which was equivalent to the number of times the increasing option had been selected over the previous ten trials. The model values options based on both the probability of reaching a given state on the next trial (s') by selecting action a (the model-based component), and on the rewards experienced in each state (the model-free component). This model is similar to other models that have assumed that subjects use state-based information to determine behavior (Daw et al., 2011; Eppinger, Walter, et al., 2013; Gläscher et al., 2010; Gureckis & Love, 2009a). Following each trial in state s and arriving in state s' after having taken action a the model computes a state prediction error (SPE):

$$\delta_{SPE} = 1 - T(s, a, s') \quad (3)$$

Next, the model updates the state transition probability:

$$T(s, a, s') = T(s, a, s') + \eta \delta_{SPE} \quad (4)$$

Here η is a free parameter that controls the learning rate for the state-transition probabilities. The state-transition probabilities for all other states not arrived at (denoted as s'') are reduced according to:

$$T(s, a, s'') = T(s, a, s') \cdot (1 - \eta) \quad (5)$$

This ensures that all transition probabilities at a given state sum to 1.

The model also tracks the model-free expected reward values for each action in each state ($Q_{MF}(s, a)$) using a SARSA learner (State-Action-Reward-State-Action) (Gläscher et al., 2010; Morris, Nevet, Arkadir, Vaadia, & Bergman, 2006). On each trial the model computes the reward prediction error (δ_{RPE}):

$$\delta_{RPE} = r - Q_{MF}(s, a) \quad (6)$$

The prediction error is then used to update the expected value for the current state action pair:

$$Q_{MF}(s, a) = Q_{MF}(s, a) + \alpha \delta_{RPE} \quad (7)$$

Here α is a free parameter that represents the learning rate for state-action pairs on each trial. The model also has the ability to allow reward information gained for actions in a specific state to be generalized across states which has been shown to improve model fits in the same task (Gureckis & Love, 2009a). For each state other than the state on the current trial (denoted as s^*) the Q_{MF} value for the same action selected on the current trial is updated:

$$Q_{MF}(s^*, a) = Q_{MF}(s^*, a) + \theta(5 * r - Q_{MF}(s^*, a)) \quad (8)$$

Here θ represents the degree to which the rewards received on each trial are generalized to the same action in different states.

After updating state-transition probabilities and expected reward value information the model then computes a *model-based* value for each action in each state ($Q_{MB}(s, a)$) using a FORWARD learner that incorporates the state-transition probabilities and the Bellman equation to determine the future value of each action (Eppinger, Walter, et al., 2013; Gläscher et al., 2010). In this task there are three possible states that participants will transition to on the next trial (s') following action a on the current trial (they can stay in the same state or move up or down one state). We estimated the Q_{MB} value for each state-action pair by the following equation:

$$Q_{MB}(s, a) = \sum_{s' \in \{s-1, s, s+1\}} T(s, a, s') * \max [Q_{MB}(s', a')] \quad (9)$$

This function multiplies the probability of transitioning to each possible state on the next trial, having taken action a in trial t , by the maximum expected reward in state s' for either action.

The model then determines a net value for each action ($Q_{Net}(s,a)$) by taking a weighted average of the model-based and model-free expected values:

$$Q_{Net}(s,a) = \omega \cdot Q_{MB}(s,a) + (1-\omega) \cdot Q_{MF}(s,a) \quad (10)$$

Where ω is a free parameter that determines the degree to which choices are based on the model-based versus model free components of the model.

Finally, the probability of selecting each action is determined using the Softmax rule:

$$P(a,t) = \frac{e^{\beta \cdot [Q_{Net}(s,a) + \pi \cdot rep(a)]}}{\sum_{j=1}^n e^{\beta \cdot [Q_{Net}(s,j) + \pi \cdot rep(j)]}} \quad (11)$$

Here β is an inverse temperature parameter that determines the degree to which participants exploit the option with the highest expected value. Larger β estimates are indicative of more consistently selecting the highest valued option, and as β approaches 0 each option is selected randomly. The autocorrelation, or perseveration parameter, π accounts for tendencies to perseverate ($\pi > 0$) or switch ($\pi < 0$) regardless of the outcome on the last trial. For the option that was selected on the prior trial $rep(a)$ is set to 1, and for all other options $rep(a) = 0$ (Daw et al., 2011; Eppinger, Walter, et al., 2013; Lau & Glimcher, 2005). In total the RL model included six free parameters: η , α , θ , ω , β , and π .

Finally, the Baseline model had one free parameter for the two choice task which represents the probability of selecting option a . This parameter is subtracted from 1 to determine the probability of selecting the other option. For the four-choice task the Baseline model had three free parameters representing the probability of selecting three of the four options on any given trial. The probability of selecting the fourth option is 1 minus the sum of the probabilities of the three other options.

Modeling Results—We fit each participant's data individually with the WSLS, RL, and Baseline models detailed above. The models were fit to the choice data from each trial by maximizing log-likelihood. We used Akaike weights to compare the relative fit of each model (Akaike, 1974; Wagenmakers & Farrell, 2004). Akaike weights are derived from Akaike's Information Criterion (AIC), which is used to compare models with different numbers of free parameters. AIC penalizes models with more free parameters. For each model, i , AIC is defined as:

$$AIC_i = -2 \log L_i + 2V_i \quad (12)$$

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 first computed AIC values for each model and for each participant's data. Akaike weights were then calculated

to obtain a continuous measure of goodness-of-fit. A difference score is computed by subtracting the AIC of the best fitting model for each data set from the AIC of each model for the same data set:

$$\Delta_i(AIC) = AIC_i - \min AIC \quad (13)$$

From the differences in AIC we then computed the relative likelihood, L , of each model, i , with the transform:

$$L(M_i | data) \propto \exp \left\{ -\frac{1}{2} \Delta_i(AIC) \right\} \quad (14)$$

Finally, the relative model likelihoods are normalized by dividing the likelihood for each model by the sum of the likelihoods for all models. This yields Akaike weights:

$$w_i(AIC) = \frac{\exp \left\{ -\frac{1}{2} \Delta_i(AIC) \right\}}{\sum_k \exp \left\{ -\frac{1}{2} \Delta_k(AIC) \right\}} \quad (15)$$

These weights can be interpreted as the probability that the model is the best model given the data set and the set of candidate models (Wagenmakers & Farrell, 2004).

We computed the Akaike weights for each model for each participant. Table 1 shows the average Akaike weights for participants in each condition. Akaike weights were highest for the RL model for both younger and older adults in every condition, except for the two-choice Increasing optimal task condition where older adults' data were best fit by the WSLs model. Akaike weights for the WSLs model were significantly higher for older than for younger adults in this condition, $t(37)=2.15$, $p<.05$, while Akaike weights for the RL model were significantly lower for older adults compared to younger adults, $t(37)=-2.11$, $p<.05$. This suggests that one potential reason why older adults performed as well as younger adults in this task was because they were using a WSLs strategy, whereas younger adults may have been using a more sophisticated model-based RL strategy.

To examine our hypothesis that younger adults would give greater weight to the model-based system than older adults we examined the best-fitting values for the ω parameter, which estimated the weight given to the model-based component of the model. A 2 (Age) X 2 (Task type) X 2 (Choice options) ANOVA revealed a significant effect of age, $F(1,174)=9.13$, $p<.01$, partial $\eta^2=.05$, where older adults' data ($M=.68$, $SD=.32$) were best fit by lower ω parameter values than data from younger adults ($M=.82$, $SD=.27$). There was also a significant main effect of task-type, $F(1,174)=5.16$, $p<.05$, partial $\eta^2=.03$, and a significant task type X choice-option interaction, $F(1,174)=5.58$, $p<.05$, partial $\eta^2=.03$. Within the Increasing-optimal task ω parameter values were significantly higher for the four-choice variant than for the two choice variant, $t(84)=-2.20$, $p<.05$, but there was no effect of choice options within the Decreasing optimal task, $t(94)=1.04$, $p=.30$. No other effects or interaction were significant.

We also examined the correlation between estimated ω parameter values and the proportion of trials participants selected the Increasing option over the course of the task. There was a strong positive association in both the Increasing-optimal ($r=.63, p<.001$) and the Decreasing-optimal task ($r=.55, p<.001$).

Association with Neuropsychological Measures

Finally, we examined whether any of the scores from the neuropsychological measures that we collected from our older adult participants were correlated with the proportion of increasing option selections and with the average ω parameter estimates that indicated the extent to which participants were using a model-based decision-making strategy. There were no significant associations between any of the neuropsychological measures and the proportion of increasing option selections. However, we observed strong positive associations between ω parameter estimates and CVLT immediate cued recall ($r=.32, p<.01$), delayed free recall ($r=.34, p<.01$), delayed cued recall ($r=.40, p<.001$), and recognition for true positives ($r=.44, p<.001$) which suggests that higher scores on these measures were associated with greater weight to the model-based component of the model.

Discussion

We examined whether older and younger adults would be more motivated to maximize the rewards they receive following each action or whether they would be motivated to take actions that improved their future state. These motivational tendencies reflect distinct decision-making strategies and are mediated by separate neural systems. Our results provide strong support for our hypothesis that older adults rely less on a goal-directed model-based system compared to younger adults. Older adults selected the option that maximized immediate reward more often than younger adults in three of the four conditions. In the only condition where they selected the Increasing option at a greater rate than younger adults, model-fits suggested that their data were better fit by a WSLS heuristic-based strategy and that they were not utilizing a model-based strategy at a greater rate than younger adults.

The examination of the best-fitting ω parameter values, which weighed the output of the model-based versus model-free components of the RL model, showed that these parameter estimates were consistently higher for younger adults than for older adults. This offers support for the hypothesis that younger adults are more likely to utilize complex model-based strategies than older adults and is in line with the findings of Eppinger and colleagues where they showed that older adults used model-based strategies to a lesser extent than younger adults in a two-stage Markov task (Eppinger, Walter, et al., 2013). This parameter was also strongly associated with the proportion of trials participants selected the increasing option, which led to improvements in the participant's state on future trials.

Our analysis of the association between the neuropsychological measures we administered to our older adult participants and the estimated ω parameter values showed robust positive associations between four measures from the CVLT and the weight given to the model-based component of the model. This is in line with work that suggests that model-based decision-making is reliant on working memory and executive function (Gershman et al., 2012; Otto, Gershman, et al., 2013; Otto, Raio, et al., 2013). It also aligns with work that

suggests that the model-based system is reliant on DLPFC (Smittenaar et al., 2013), an area that plays a critical role in working memory (Curtis & D'Esposito, 2003). However, it is important to note that Eppinger and colleagues found a link between working memory and model-based decision-making for younger adults, but not for older adults (Eppinger, Walter, et al., 2013). This discrepancy could be due to differences in the decision-making tasks and measures of working memory between the two studies. Additionally, it's also important to note that we did not observe any link between model-based responding and other measures of working memory or executive attention aside from the four measures from the CVLT, which might have been expected. However, we did observe weak but positive associations between best-fitting ω parameter values and performance on the Digit Span ($r=.16, p=.10$), and the FAS ($r=.19, p=.06$). Future work should further examine the link between model-based learning and measures of working memory and executive function in both younger and older adults. An additional question is whether reduced model-based responding predicts more severe age-related cognitive decline.

It is important to note that in the Decreasing optimal tasks younger adults showed a strong tendency to value improving their state over maximizing reward, even when doing so was disadvantageous. Younger adults even selected the sub-optimal Increasing option more frequently as the task progressed, while older adults selected the Increasing option at the same rate across the four-choice variant of the task and at a slightly higher rate toward the end of the two-choice variant of the task. This suggests that younger adults are less likely to be satisfied with simply obtaining greater reward and are instead motivated by the goal of improving their future state, even when doing so is counterproductive. The fact that younger adults showed a strong tendency to select the poorer, increasing option more over course of the task might be surprising in that one could reasonably assume that if participants understood the contingencies of the task and sufficiently engaged then we would observe a higher proportion of optimal choices by the end of the task. We believe our results support our assertion that younger adults are more motivated to improve their future state than to maximize reward, and that the higher proportions of increasing option selections in both tasks is due to their learning the state-transition probabilities over time and selecting the Increasing option which has a higher probability of improving their future state.

Additionally, it is also important to note that younger adults selected the increasing option more frequently in the four-choice variant of the Increasing-optimal task than in the two-choice variant. While speculative, one possibility is that younger adults viewed the four-choice task as more challenging and were more motivated to take actions that improved their future state. Younger adults, who we argue have a motivational bias toward improving their future state, may have felt more motivated by a more challenging four-choice task and this may have enhanced their default motivation to attempt to improve their state. However, we acknowledge that we did not predict, a priori, an advantage for younger adults performing the four-choice task over those performing the two-choice task. Future work should further examine this issue.

Overall, the results from the two-choice tasks in the current study replicate those from our previous paper (Worthy et al., 2011). However, one subtle difference is that in the previous paper older adults selected the increasing option less over time in the Decreasing-optimal

task, but in the current paper older adults selected the increasing option slightly more over time. However, a repeated measures ANOVA showed no interaction between block and the study that the data came from, $F(4,192)=1.20, p=.31$. Thus, the data from the present study are consistent with our previously published paper (Worthy et al., 2011).

The results of our study are also consistent with a previous meta-analysis by Mata and colleagues that suggests that older adults tend to stick with behaviors that are initially beneficial to a greater extent than younger adults, but they must learn to avoid these options if they prove detrimental over time (Mata, Josef, Samanez-Larkin, & Hertwig, 2011). In our tasks the decreasing option may appear more initially rewarding than the increasing option in the Decreasing-optimal task, where it provides 60 more points on each trial, than in the Increasing-optimal task, where it provides only 10 more points on each trial. Because of this, older adults may have a stronger tendency to stick with the decreasing option in the Decreasing-optimal task, but they may be more willing to engage in heuristic-based WSL strategies in the Increasing optimal task in order to improve their rewards on future trials. Additionally, this interpretation is also consistent with the association between memory measures and model-based decision-making in older adults where older adults with greater memory ability may be more capable of overcoming the initial bias toward model-free strategy use in favor of model-based strategy use.

One important question regarding our results is whether they suggest that older adults are more motivated to pursue immediate rewards than to improve their future state, per se, or whether they lack sufficient cognitive resources to engage in the model-based decision-making that is required to improve one's future state. In other words, are our findings due to a difference in motivation or in cognitive resources? It's very likely that older adults would want to improve their future state if they had knowledge that it was productive to do so, but in our tasks they may have been unable to assess the degree to which each action improved their future state due to declines in cognitive resources. The association between model-based strategy-use and four measures of the CVLT support the assertion that our results may have been due to differences in cognitive ability than in motivational orientation. However, one thing that seems clear, particularly from the Decreasing-optimal task results, is that younger adults are more motivated than older adults to improve their future state rather than maximize immediate reward even when doing so is counter-productive. Presumably younger adults have the cognitive resources necessary to engage in either model-free or model-based decision-making and they appear inclined to engage in model-based decision-making even when doing so is sub-optimal.

An additional point to note is that there is much overlap between the neural circuits that mediate motivation and cognition which make it difficult to determine the extent to which our results are uniquely attributable to motivation versus cognitive ability (Berridge & Robinson, 2003). It is also entirely possible that cognitive and motivational processes interact. For example, people may be less motivated to try to improve their future state if they do not have the necessary cognitive capacity to do so. Given declines in cognitive resources older adults may reorient their decision-making goals and strategies to suit what they are more capable of doing, like maximizing immediate reward rather than attempting to improve their future state. However, we acknowledge that this hypothesis is speculative, and

focusing on the degree to which age related differences in decision-making are due to cognitive versus motivational differences is a key direction for future work. One possible avenue might be to incorporate the use of alternative methods like collecting self-report data that might better assess participants' motivation and goals during decision-making.

Another key direction for future work is to more precisely identify the links between age-related neurobiological changes in striatal and prefrontal areas that are critical for model-free and model-based decision-making. Aging has been associated with declines in these regions (Raz et al., 2005; West, 1996) and with reduced integrity of the dopaminergic system that is thought to be critical for both model-based and model-free reinforcement learning (Bäckman, Nyberg, Lindenberg, Li, & Farde, 2006; Li, Lindenberg, & Sikström, 2001). Additionally, older adults have shown reduced striatal responses to reward prediction errors (Chowdhury et al., 2013; Eppinger, Schuck, Nystrom, & Cohen, 2013). This could affect both model-based and model-free decision-making as prediction errors for both reward and state-based representations are tracked by the striatum (Daw et al., 2011). One possibility is that age-related declines in striatal regions are at least partially responsible for the older adults' reduced ability to update state-transition probabilities, which are critical for model-based decision-making. Declines in prefrontal regions like DLPFC may also hurt model-based decision-making as this regions may be critical for computing model-based expected values which incorporate both state-transition probabilities and model-free reward values (Smittenaar et al., 2013). Our analysis of the neuropsychological memory measures suggest that preserved memory ability may be linked to preserved model-based decision-making.

One important theoretical question that our study only partially addresses is how heuristic-based WSLs strategies fit in model-free versus model-based RL strategies. Prior work suggests that older adults show stronger tendencies than younger adults to use heuristics (Castel, Rossi, & McGillivray, 2012; Worthy & Maddox, 2012). In the current work we only found strong evidence for older adults utilizing a heuristic-based WSLs strategy in the two-choice variant of the Increasing optimal task. One possibility, which we admit is speculative, is that in decision-making situations older adults have a default tendency to utilize a WSLs heuristic to choose amongst alternatives. If the WSLs strategy is perceived to be working well then it is utilized further, but if it appears inadequate, as in the Decreasing-optimal task where the increasing option provides much smaller rewards than the decreasing option, then older adults may abandon it and switch to a model-free RL strategy. Future work should be done using tasks that might more precisely distinguish between RL and heuristic-based strategies.

Conclusion

Motivation is central to action and it can shape how we think about and approach decision-making situations. A useful avenue for research on the motivation-cognition interface is to examine differences between motivation that is state-based versus reward-based. State-based motivation is goal-directed and centered on developing a model of the environment and taking actions that improve future states. Reward-based motivation is habitual and centered on taking actions that directly lead to reward. Our results suggest that aging may shift the

balance between these two motivational foci to where older adults are motivated more by the habitual tendency to obtain reward and younger adults are motivated by the goal-directed tendency to improve their state. The association between model-based strategy use and the memory measures from our neuropsychological test battery suggest that this shift may be due to age-related declines in fluid intelligence (Salthouse, 2004) that prevent older adults from engaging in more computationally demanding model-based decision-making. Future work is needed to more precisely delineate factors that influence state-based and reward-based motivation including work aimed at dissociating the distinct neural processes involved.

Acknowledgments

This research was funded by NIA grant AG043425 to DAW and WTM and NIDA grant DA032457 to WTM. We thank Anna Anthony for help with data collection.

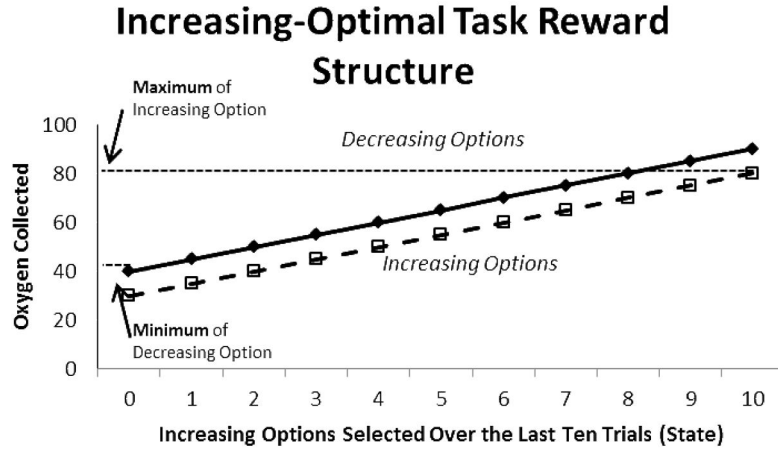
References

- Akaike H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*. 1974; 19(6):716–723.
- Bäckman L, Nyberg L, Lindenberger U, Li SC, Farde L. The correlative triad among aging, dopamine, and cognition: current status and future prospects. *Neuroscience & Biobehavioral Reviews*. 2006; 30(6):791–807. [PubMed: 16901542]
- Berridge KC, Robinson TE. Parsing reward. *Trends in neurosciences*. 2003; 26(9):507–513. [PubMed: 12948663]
- Byrne KA, Worthy DA. Do narcissists make better decisions? An investigation of narcissism and dynamic decision-making performance. *Personality and Individual Differences*. 2013; 55(2):112–117.
- Castel AD, Rossi AD, McGillivray S. Beliefs about the “hot hand” in basketball across the adult life span. *Psychology and aging*. 2012; 27(3):601. [PubMed: 22288426]
- Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, Dolan RJ. Dopamine restores reward prediction errors in old age. *Nature neuroscience*. 2013; 16(5):648–653.
- Cooper JA, Worthy DA, Gorlick MA, Maddox WT. Scaffolding Across the Lifespan in History-Dependent Decision Making. *Psychology and Aging*. 2013; 28(2):505–514. [PubMed: 23795765]
- Curtis CE, D’Esposito M. Persistent activity in the prefrontal cortex during working memory. *Trends in cognitive sciences*. 2003; 7(9):415–423. [PubMed: 12963473]
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans’ choices and striatal prediction errors. *Neuron*. 2011; 69(6):1204–1215. [PubMed: 21435563]
- Denburg NL, Tranel D, Bechara A. The ability to decide advantageously declines prematurely in some normal older persons. *Neuropsychologia*. 2005; 43(7):1099–1106. [PubMed: 15769495]
- Denburg NL, Weller JA, Yamada TH, Shivapour DM, Kaup AR, LaLoggia A, Bechara A. Poor decision making among older adults is related to elevated levels of neuroticism. *Annals of Behavioral Medicine*. 2009; 37(2):164–172. [PubMed: 19350336]
- Eppinger B, Schuck NW, Nystrom LE, Cohen JD. Reduced striatal responses to reward prediction errors in older compared with younger adults. *The Journal of Neuroscience*. 2013; 33(24):9905–9912. [PubMed: 23761885]
- Eppinger B, Walter M, Heekeren HR, Li S-C. Of goals and habits: age-related and individual differences in goal-directed decision-making. *Frontiers in neuroscience*. 2013;7. [PubMed: 23378827]
- Fridlund, A.; Delis, DC. *CVLT research edition administration and scoring software*. New York: The Psychological Corporation; 1987.
- Gershman, SJ.; Markman, AB.; Otto, AR. *Retrospective Revaluation in Sequential Decision Making: A Tale of Two Systems*. 2012.

- Gläscher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010; 66(4):585–595. [PubMed: 20510862]
- Gureckis TM, Love BC. Learning in noise: Dynamic decision-making in a variable environment. *Journal of Mathematical Psychology*. 2009a; 53(3):180–193. [PubMed: 20161328]
- Gureckis TM, Love BC. Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition*. 2009b; 113(3):293–313. [PubMed: 19427635]
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience*. 2008; 28(22):5623–5630. [PubMed: 18509023]
- Heaton, RK. A manual for the Wisconsin card sorting test. Western Psychological Services; 1981.
- Lau B, Glimcher PW. Dynamic response-by-response models of matching behavior in rhesus monkeys. *Journal of the experimental analysis of behavior*. 2005; 84(3):555–579. [PubMed: 16596980]
- Lezak, M. Neuropsychological testing. University Press; Oxford: 1995.
- Li SC, Lindenberger U, Sikström S. Aging cognition: from neuromodulation to representation. *Trends in cognitive sciences*. 2001; 5(11):479–486. [PubMed: 11684480]
- Maddox WT, Gorlick MA, Worthy DA, Beevers CG. Depressive symptoms enhance loss-minimization, but attenuate gain-maximization in history-dependent decision-making. *Cognition*. 2012
- Mata R, Josef AK, Samanez-Larkin GR, Hertwig R. Age differences in risky choice: a meta-analysis. *Annals of the New York Academy of Sciences*. 2011; 1235(1):18–29. [PubMed: 22023565]
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. Midbrain dopamine neurons encode decisions for future action. *Nature neuroscience*. 2006; 9(8):1057–1063.
- O'Doherty JP. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology*. 2004; 14(6):769–776. [PubMed: 15582382]
- Otto AR, Gershman SJ, Markman AB, Daw ND. The Curse of Planning Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive. *Psychological science*. 2013; 24(5):751–761. [PubMed: 23558545]
- Otto AR, Gureckis TM, Markman AB, Love BC. Navigating through abstract decision spaces: Evaluating the role of state generalization in a dynamic decision-making task. *Psychonomic Bulletin & Review*. 2009; 16(5):957–963. [PubMed: 19815805]
- Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*. 2013; 110(52):20941–20946.
- Otto AR, Taylor EG, Markman AB. There are at least two kinds of probability matching: Evidence from a secondary task. *Cognition*. 2011; 118(2):274–279. [PubMed: 21145046]
- Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Acker JD. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebral cortex*. 2005; 15(11):1676–1689. [PubMed: 15703252]
- Reed AE, Mikels JA, Simon KI. Older adults prefer less choice than young adults. *Psychology and aging*. 2008; 23(3):671. [PubMed: 18808256]
- Salthouse TA. What and when of cognitive aging. *Current Directions in Psychological Science*. 2004; 13(4):140–144.
- Salthouse TA. When does age-related cognitive decline begin? *Neurobiology of Aging*. 2009; 30(4):507–514. [PubMed: 19231028]
- Samanez-Larkin GR, Kuhnen CM, Yoo DJ, Knutson B. Variability in nucleus accumbens activity mediates age-related suboptimal financial risk taking. *The Journal of Neuroscience*. 2010; 30(4):1426–1434. [PubMed: 20107069]
- Schwartz B. Self-determination: The tyranny of freedom. *American Psychologist*. 2000; 55(1):79. [PubMed: 11392868]
- Schwartz, B. The paradox of choice. HarperCollins; 2009.

- Smittenaar P, FitzGerald THB, Romei V, Wright ND, Dolan RJ. Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron*. 2013; 80(4):914–919. [PubMed: 24206669]
- Sutton, RS.; Barto, AG. Reinforcement learning: An introduction. Vol. 1. Cambridge Univ Press; 1998.
- Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. *Psychonomic bulletin & review*. 2004; 11(1):192–196. [PubMed: 15117008]
- Wechsler, D. WAIS-III, Wechsler Adult Intelligence Scale: Administration and Scoring Manual. Psychological Corporation; 1997.
- West RL. An application of prefrontal cortex function theory to cognitive aging. *Psychological bulletin*. 1996; 120(2):272. [PubMed: 8831298]
- Worthy DA, Gorlick MA, Pacheco JL, Schnyer DM, Maddox WT. With Age Comes Wisdom Decision Making in Younger and Older Adults. *Psychological Science*. 2011; 22(11):1375–1380. [PubMed: 21960248]
- Worthy DA, Hawthorne MJ, Otto AR. Heterogeneity of strategy use in the Iowa gambling task: A comparison of win-stay/lose-shift and reinforcement learning models. *Psychonomic bulletin & review*. 2013; 20:364–371. [PubMed: 23065763]
- Worthy DA, Maddox WT. Age-based differences in strategy use in choice tasks. *Frontiers in neuroscience*. 2012;5. [PubMed: 22303292]
- Worthy DA, Otto AR, Maddox WT. Working-Memory Load and Temporal Myopia in Dynamic Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2012; 38(6):1640–1658.
- Yechiam E, Busemeyer JR. Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic bulletin & review*. 2005; 12(3):387–402. [PubMed: 16235624]

a.



b.

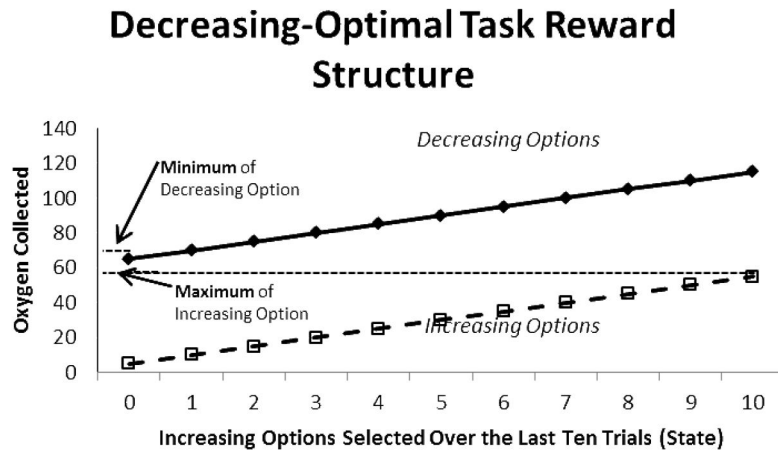


Figure 1.

(a.) Reward structure for Increasing-optimal task. The amount of oxygen provided when each option was selected depended on the number of times participants had selected from the increasing options over the past ten trials. In the two-choice task there was one increasing and one decreasing option. In the four-choice task there were two of each type of option. The optimal strategy was to repeatedly select the increasing option. (b.) Reward structure for Decreasing-optimal task. The optimal strategy was to repeatedly select the decreasing option because the maximum that could be given by the increasing option was less than the minimum given by repeatedly selecting the decreasing option.

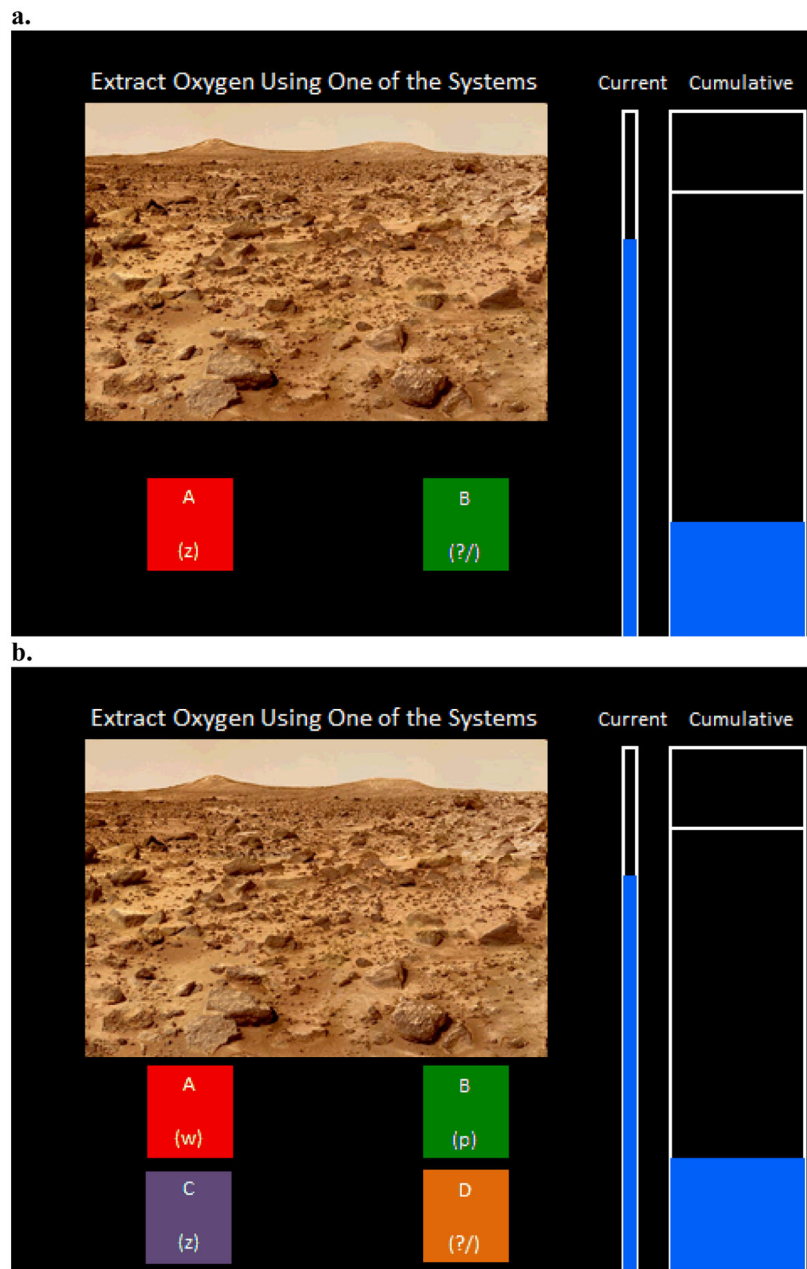
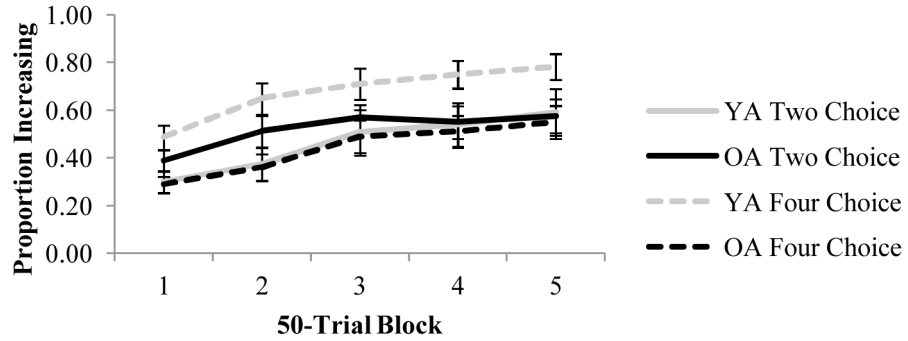


Figure 2.

(a.) Sample screen shot for the two-choice tasks. On each trial participants would select one of the systems to use to extract oxygen. The amount of oxygen extracted on each trial would be indicated in the “Current” tank and then transferred to the “Cumulative” tank. (b.) Sample screen shot for the four-choice tasks.

a.

Increasing Optimal Task: Proportion of Increasing Option Selections



b.

Decreasing Optimal Task: Proportion of Increasing Option Selections

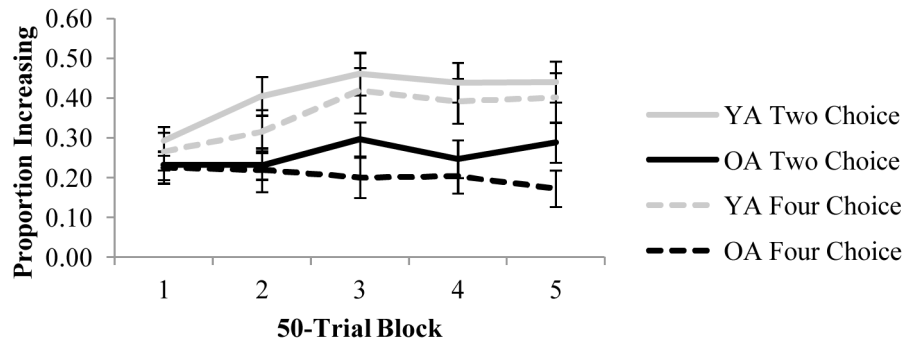


Figure 3. Proportion of times participants selected the increasing option in each task in 50-trial blocks. The increasing option was the optimal choice in the Increasing optimal task (a), but the sub-optimal choice in the Decreasing-optimal task (b). Error bars represent standard errors of the mean.

Table 1

Z-scores Summary for each Neuropsychological Exam for Participants in the Increasing Optimal Task

Neuropsychological Test	Two-Choice		Four-Choice	
	Mean (SD)	Range	Mean (SD)	Range
Digit Span	0.16 (0.86)	-1.3-2.3	0.76 (1.11)	-1.0-3.0
CVLT Delayed Recall (Free)	0.50 (0.86)	-1.0-2.0	0.98 (0.92)	-1.0-2.5
CVLT Immediate Recall (Free)	0.77 (0.74)	-0.5-2.0	0.96 (1.14)	-2.0-2.5
CVLT Delayed Recall (Cued)	0.43 (0.82)	-1.0-2.0	0.76 (0.79)	-0.5-2.0
CVLT Immediate Recall (Cued)	0.50 (0.77)	-0.5-1.5	0.78 (0.83)	-1.0-2.5
CVLT Recognition False Positives	-0.11 (1.23)	-1.0-3.0	-0.28 (0.88)	-1.0-2.0
CVLT Recognition True Positives	-0.09 (1.00)	-1.0-3.0	0.28 (0.93)	-2.5-1.0
FAS	-0.34 (1.17)	-3.0-1.6	0.22 (1.15)	-1.8-2.5
Trails A	-0.24 (0.88)	-1.3-1.9	-0.74 (0.53)	-1.4-0.8
Trails B	-0.31 (0.94)	-1.1-1.2	-0.71 (0.28)	-1.2-0.1
WCST Errors	0.41 (0.96)	-1.5-2.1	0.63 (0.60)	-0.4-1.8
WCST Perseveration	0.69 (1.00)	-0.8-3.0	0.43 (0.43)	-0.3-1.3
Demographic Information	Two-Choice		Four-Choice	
Age	67.59 (6.31)	60-88	67.66 (5.04)	61-81
Years of Education	16.09 (2.65)	10-20	18.04 (2.24)	16-24

Note: Mean z-scores for each exam with standard deviation in parenthesis and z-score range. Scores are separated by condition (two-choice or four-choice).

Table 2

Z-scores Summary for each Neuropsychological Exam for Participants in the Decreasing Optimal Task

Neuropsychological Test	Two-Choice		Four-Choice	
	Mean (SD)	Range	Mean (SD)	Range
Digit Span	0.17 (0.74)	-1.3-1.7	0.31 (0.59)	-0.3-2.0
CVLT Delayed Recall (Free)	0.37 (1.12)	-3.0-2.0	0.55 (0.94)	-1.0-2.5
CVLT Immediate Recall (Free)	0.89 (1.06)	-1.5-2.5	0.65 (0.81)	-0.5-2.0
CVLT Delayed Recall (Cued)	0.74 (0.88)	-2.0-2.0	0.55 (0.84)	-1.0-2.0
CVLT Immediate Recall (Cued)	0.87 (1.07)	-2.5-2.5	0.70 (0.78)	-1.0-2.5
CVLT Recognition False Positives	-0.54 (0.54)	-1.0-1.0	-0.40 (0.80)	-1.0-2.5
CVLT Recognition True Positives	-0.17 (0.91)	-2.5-1.0	0.05 (0.83)	-2.0-1.0
FAS	-0.21 (0.92)	-1.8-2.1	0.32 (1.07)	-1.2-2.5
Trails A	-0.15 (0.88)	-1.7-1.6	-0.52 (0.51)	-1.4-0.9
Trails B	-0.18 (0.94)	-1.5-3.0	-0.66 (0.55)	-2.1-0.5
WCST Errors	0.83 (1.52)	-2.3-2.5	0.32 (0.91)	-1.6-2.5
WCST Perseveration	0.87 (1.19)	-1.9-2.5	0.38 (0.84)	-1.3-2.5
Demographic Information	Two-Choice		Four-Choice	
Age	67.65 (5.30)	61-78	67.25 (6.28)	61-83
Years of Education	17.67 (4.29)	12-27	17.75 (1.68)	13-20

Note: Mean z-scores for each exam with standard deviation in parenthesis and z-score range. Scores are separated by condition (two-choice or four-choice).

Table 3

Akaike Weights for Each Model

	WSLS	RL	Baseline
Increasing Optimal Task			
<i>Two-choice</i>			
Younger Adults	.28 (.43)	.72 (.43)	.01 (.03)
Older Adults	.59 (.47)	.41 (.48)	.00 (.00)
<i>Four-choice</i>			
Younger Adults	.07 (.22)	.81 (.39)	.11 (.31)
Older Adults	.17 (.36)	.59 (.48)	.24 (.43)
Decreasing Optimal Task			
<i>Two-choice</i>			
Younger Adults	.30 (.37)	.67 (.37)	.03 (.10)
Older Adults	.09 (.23)	.79 (.35)	.30 (.06)
<i>Four-choice</i>			
Younger Adults	.07 (.22)	.85 (.30)	.07 (.23)
Older Adults	.17 (.36)	.73 (.43)	.09 (.30)

Note: Standard deviations are listed in parentheses

Table 4

Average Best-Fitting Parameter Estimates for the RL Model

	Increasing Optimal Task			
	<i>Two-choice</i>		<i>Four-choice</i>	
	Younger	Older	Younger	Older
State learning rate (η)	.69 (.38)	.67 (.39)	.39 (.38)	.45 (.42)
Reward learning rate (α)	.25 (.29)	.51 (.38)	.33 (.39)	.65 (.41)
Reward generalization rate (θ)	.15 (.26)	.03 (.07)	.19 (.31)	.29 (.33)
Model-based weight (ω)	.62 (.36)	.61 (.31)	.86 (.25)	.68 (.32)
Inverse temperature (β)	1.84 (2.12)	.78 (1.30)	2.17 (2.01)	1.70 (1.98)
Perseveration (π)	4.12 (5.89)	4.54 (5.25)	.09 (4.48)	1.78 (5.37)
	Decreasing Optimal Task			
	<i>Two-choice</i>		<i>Four-choice</i>	
	Younger	Older	Younger	Older
State learning rate (η)	.52 (.38)	.22 (.33)	.30 (.38)	.64 (.41)
Reward learning rate (α)	.33 (.44)	.38 (.42)	.42 (.43)	.25 (.39)
Reward generalization rate (θ)	.28 (.42)	.29 (.43)	.40 (.48)	.17 (.32)
Model-based weight (ω)	.87 (.19)	.77 (.32)	.88 (.22)	.65 (.34)
Inverse temperature (β)	1.10 (1.57)	2.72 (2.36)	2.25 (2.16)	2.67 (2.27)
Perseveration (π)	3.51 (6.81)	2.95 (5.74)	.84 (3.52)	3.49 (5.99)

Note: Standard deviations are listed in parentheses