Cognitive Psychology 83 (2015) 40–53

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Chronic motivational state interacts with task reward structure in dynamic decision-making

Jessica A. Cooper [a,*], Darrell A. Worthy [b], W. Todd Maddox [a]

[a] The University of Texas at Austin, Department of Psychology, 108 E. Dean Keeton, A8000, Austin, TX 78712, United States
[b] Texas A&M University, Department of Psychology, 4235 TAMU, College Station, TX 77843, United States

ARTICLE INFO

Article history:
Accepted 26 September 2015
Available online 29 October 2015

Keywords:
Motivation
Decision-making
Regulatory fit
Regulatory focus
Reward

ABSTRACT

Research distinguishes between a habitual, model-free system motivated toward immediately rewarding actions, and a goal-directed, model-based system motivated toward actions that improve future state. We examined the balance of processing in these two systems during state-based decision-making. We tested a regulatory fit hypothesis (Maddox & Markman, 2010) that predicts that global trait motivation affects the balance of habitual- vs. goal-directed processing but only through its interaction with the task framing as gain-maximization or loss-minimization. We found support for the hypothesis that a match between an individual's chronic motivational state and the task framing enhances goal-directed processing, and thus state-based decision-making. Specifically, chronic promotion-focused individuals under gain-maximization and chronic prevention-focused individuals under loss-minimization both showed enhanced state-based decision-making. Computational modeling indicates that individuals in a match between global chronic motivational state and local task reward structure engaged more goal-directed processing, whereas those in a mismatch engaged more habitual processing.

* Corresponding author at: Department of Psychology, The University of Texas, 108 E. Dean Keeton, Stop A8000, Austin, TX 78712, United States. Fax: +1 (512) 471 6175.

E-mail addresses: jessica.cooper@utexas.edu (J.A. Cooper), worthyda@tamu.edu (D.A. Worthy), maddox@psy.utexas.edu (W.T. Maddox).

http://dx.doi.org/10.1016/j.cogpsych.2015.09.001
0010-0285/

1. Introduction

Motivation is a key feature of decision-making that is often studied in terms of approaching positive states and avoiding negative states (e.g. Atkinson, 1964; Bandura, 1986; Roseman, Spindel, & Jose, 1990). We, along with others (e.g., Braver et al., 2014; Maddox & Markman, 2010), argue that the most common definition of motivation as a simple increase in effortful cognitive processing (i.e., trying harder) is outdated, and that a deeper understanding of the complex motivation–cognition interface is crucial to theorizing about motivation as well as cognition (see Braver et al., 2014, for a review).

Under more recent views, motivation is thought to operate at multiple levels, and the effects of motivation on behavior derive from the interactions between these levels. The interactive nature of motivation on behavior is captured by the notion of "regulatory fit" (Higgins, 2000; Maddox & Markman, 2010), which is achieved when the individual's global motivational state (chronic or situational) aligns with the local motivational task framing. Importantly, approach or avoidance motivation at one level can have vastly different effects on behavior depending upon the valence of motivation at another level. To date, little work has explored these multi-level motivational effects on the balance of cognitive processing. This is the focus of the present report.

Regulatory fit effects have been shown in a variety of domains including judgments of morality (Camacho, Higgins, & Luger, 2003), communication effectiveness (Aaker & Lee, 2001; Cesario, Grant, & Higgins, 2004), and generation of anagram solutions (Shah, Higgins, & Friedman, 1998). Unfortunately, no strong mechanistic explanations for these regulatory fit effects have been offered, mainly because these tasks are ones for which no unique optimal strategy can be defined. This shortcoming has been addressed by examining tasks for which the optimal strategy is uniquely identifiable, and importantly is mediated by a specific cognitive process. This work tests the hypothesis that a "fit" between the global and local motivational state enhances effortful cognitive processing at the expense of automatic habitual processing (see Maddox & Markman, 2010, for a review). Critically, whether this enhanced effortful processing leads to better performance depends upon whether optimal task performance is mediated by effortful processing. Thus, this work argues that the motivation–cognition relationship involves a three-way interaction between the global motivational state, the local motivational state, and the cognitive processing system that optimally mediates task performance. When the task is one for which optimal performance requires effortful processing, a regulatory fit is advantageous. However, when the task is one for which optimal performance requires automatic habitual processing, such as implicit category learning (Grimm, Markman, Maddox, & Baldwin, 2008), a regulatory mismatch is advantageous. Tests of this three-way interaction find support in studies that examine category learning (e.g. Grimm et al., 2008; Maddox, Baldwin, & Markman, 2006) and decision-making (Otto, Markman, Gureckis, & Love, 2010; Worthy, Maddox, & Markman, 2007).

Regulatory fit effects in decision-making have shown that decision-makers in a regulatory fit more often choose to systematically explore their environment, while those in a regulatory mismatch more often exploit the highest-valued option (Otto et al., 2010; Worthy et al., 2007). Worthy et al. (2007) and Otto et al. (2010), like much work on regulatory fit, focused on the effect of situational (or induced) regulatory focus. *Situational* or experimentally-induced motivational focus, obtained by making individuals temporarily experience either a subjective history of promotion success or prevention success (Higgins et al., 2001), is extremely helpful in providing methods for boosting overall task performance, but does little to identify performance advantages related to stable traits of the individual. *Chronic* promotion and prevention focus is measured using the Regulatory Focus Questionnaire (RFQ; Higgins et al., 2001), which provides scores for two motivational modes that affect the sensitivity of the motivational system: a promotion focus, in which one becomes more sensitive to potential gains, and a prevention focus, in which one becomes more sensitive to potential losses (Higgins, 1997). Here, we examine whether these chronic (trait-driven) dispositional tendencies lead people to engage in qualitatively different decision-making strategies depending on task framing, whether the goal of the task is to maximize gains or minimize losses. Simply altering the way that decision-making tasks are framed is predicted to drastically alter the types of decision-making strategies that people implement, and could provide a simple mechanism for optimizing performance in decision-making tasks by catering the framing to an individual's chronic motivational mode.

1.1. The current study

In the current work, we test the hypothesis that chronic (trait-driven) motivational modes interact with reward structure to affect decision-making performance and cognitive processing by evaluating decision-making performance and utilizing computational modeling to quantify the degree to which participants use goal-directed versus habitual strategies. Habitual and goal-directed systems are often referred to as model-free and model-based in the neuroscience literature (e.g. Daw, Niv, & Dayan, 2005; Doya, Samejima, Katagiri, & Kawato, 2002; Gläscher, Daw, Dayan, & O'Doherty, 2010). The model-free, or habitual, system is motivated toward actions that lead directly to reward while the model-based, or goal-directed, system is more computationally demanding and requires consideration of how actions can affect both immediate and future outcomes. The distinction that use of the goal-directed system is more computationally demanding has been reflected in several recent studies that have found a relationship between goal-directed strategies and working memory processes (e.g. Gershman, Markman, & Otto, 2014; Otto, Gershman, Markman, & Daw, 2013).

The decision-making task that we utilize has been widely used to evaluate state-based decision-making (e.g. Cooper, Worthy, Gorlick, & Maddox, 2013; Gureckis & Love, 2009a; Worthy, Cooper, Byrne, Gorlick, & Maddox, 2014). In this task participants repeatedly choose between two rewarding options and gain information about the reward environment by making decisions and receiving rewards. One option always provides a larger immediate reward but causes the rewards available on future trials to decrease. The other option provides a lower immediate reward on each trial but causes the available future rewards to increase by improving one's future state. Thus, options that favor the goal-directed system are directly pitted against those that favor the habitual, reward-based system. Importantly, this task is amenable to computational modeling that can be used to quantify strategy engagement, allowing us to directly link the balance of processing with differences in performance (Worthy et al., 2014). We fit participants' data with a series of computational models, including a model that includes a free parameter that quantifies the weight placed on the output from goal-directed or habitual processing systems.

We predict that a regulatory fit, known to shift the cognitive balance toward goal-directed processing (e.g. Markman, Maddox, Worthy, & Baldwin, 2007; Worthy et al., 2007) and to aid complex problem-solving (e.g. Grimm et al., 2008), facilitates the use of goal-directed processes. Specifically, we predict that individuals in a match between chronic regulatory-focus and reward structure will be more motivated toward improvements in state than those in a mismatch. We expect that a regulatory match will result in differences in performance and in computational model parameter estimates reflecting the increased use of more computationally-demanding, goal-directed strategies.

We test the prediction that chronic regulatory focus interacts with the task reward structure in two versions of the decision-making task. In one version participants are asked to minimize losses, while in the other they are asked to maximize gains. Importantly, the underlying reward structure, optimal strategy, and interaction between the reward options is identical between these conditions—the only difference is that all of the reward values in the loss-minimization condition are shifted by a constant so that they are negative and provide losses on each trial. We predict that in the gain-maximization condition, individuals who are more promotion-focused will experience a regulatory match and will perform better than those who are prevention-focused. Additionally, in the loss-minimization condition we predict that prevention-focused individuals will experience a regulatory match and will outperform promotion-focused individuals.

2. Methods

2.1. Participants

Participants were 80 young adults (M_{age} = 18.5, SD = .64) recruited from the University of Texas, with 20 participants in each group. Each participant was compensated $10 per hour for his or her participation. Informed consent was obtained from all participants, and the experiment was approved for ethics procedures using human participants at the University of Texas.

2.2. Materials and procedure

Students completed the Regulatory Focus Questionnaire (RFQ; Higgins et al., 2001) as part of the introductory to psychology participant screening. Prevention and promotion focus scores were calculated for each individual according to scoring instructions. A relative promotion-focus score was calculated for each participant by subtracting his or her prevention-focus score from his or her promotion-focus score. Individuals with a positive relative promotion-focus score were classified as promotion focused, while individuals with a prevention-focus score higher than their promotion-focus score (negative relative promotion-focus score) were classified as prevention focused. Promotion focused and prevention focused individuals were invited to our lab to complete the study. Upon arriving, participants were screened again with the RFQ to ensure that their scores were consistent. Participants who had equal promotion and prevention focus scores were not included in our study, as they did not fit into either group. Participants were not informed of their promotion or prevention focus status. Demographic information for each group can be observed in Table 1.

Participants completed the Mars Farming task in one of two conditions: gain-maximization or loss-minimization. The task was completed on PC computers using Psychtoolbox 2.54 for MATLAB (Brainard, 1997; Pelli, 1997). Each task consisted of 250 trials in which participants would choose between two reward options.

2.2.1. Gain-maximization

In the gain-maximization condition, participants were asked to select one of two systems for extracting oxygen on each trial. Following each choice, a small tank filled with the oxygen extracted on that trial. This was then transferred to a larger tank. A line on the larger tank corresponded to the amount of oxygen needed to sustain life on Mars, and participants were told to collect at least that much oxygen over the course of the experiment. A sample screen shot from the gain-maximization condition is displayed in Fig. 1A (left panel). Fig. 2A (top panel) shows the reward structure for the gain-maximization task. Participants were not told about the reward contingencies of the two options and had to learn about them through trial and error. The reward associated with selecting each option is dependent on the number of times that one of the options, referred to as the increasing option, has been selected over the previous 10 trials (the state). For example, if the increasing option had been selected on four of the previous 10 trials, then 35 units of oxygen would be extracted with the increasing option system, whereas 75 units of oxygen would be extracted with the other system. Choosing the increasing option causes the gain for both options to increase on future trials whereas choosing the other option, referred to as the decreasing option, causes the gain for both options to decrease on future trials. However, the decreasing option always gives a higher gain on any given trial, as suggested by the example above. Thus, on every trial the decreasing option will yield a higher gain, but the more that the increasing option is selected the larger the gain obtained for both options will be on future trials. If participants select the increasing option on each trial, they would eventually reach the highest state (10), whereas selecting the decreasing option on each trial would eventually lead to the lowest state (0). Selecting the increasing option involves a cost in the immediate gain but a delayed benefit in overall gain, whereas selecting the decreasing option involves a benefit in the immediate gain but a delayed cost in overall gain. The goal line was drawn on the oxygen collection tank (Fig. 1A). The positioning of this line corresponded to performance that would be attained by selecting the optimal choice[1] (increasing option) on approximately 80% of trials.

2.2.2. Loss-minimization

A sample screen shot from a trial in the loss-minimization condition of the Mars Farming task is displayed in Fig. 1B (right panel). Fig. 2B (bottom panel) shows the reward structure for the task under loss-minimization conditions. In the loss-minimization condition, participants were told that they would be testing two oxygen extraction systems that both extract oxygen at the same rate but that

[1] With full knowledge of task length an ideal actor would select the increasing option until the last 8–9 trials, then switch to the decreasing option. Participants were not provided information regarding trial number, and our analysis will treat the increasing option as the optimal choice on all trials.

Table 1
Demographic information.

	Gain-maximization		Loss-minimization	
	Promotion	Prevention	Promotion	Prevention
Age	18.67 (.68)	18.82 (.60)	18.38 (.87)	18.47 (.52)
Gender	10 F/10 M	14 F/6 M	12 F/8 M	14 F/6 M
RFQ difference	5.45 (2.74)	−3.80 (2.07)	5.10 (3.43)	−4.50 (1.64)

Note: Standard deviations in parenthesis. RFQ difference score is promotion focus score minus prevention focus score; positive values indicate higher promotion focus, and negative indicate higher prevention focus.

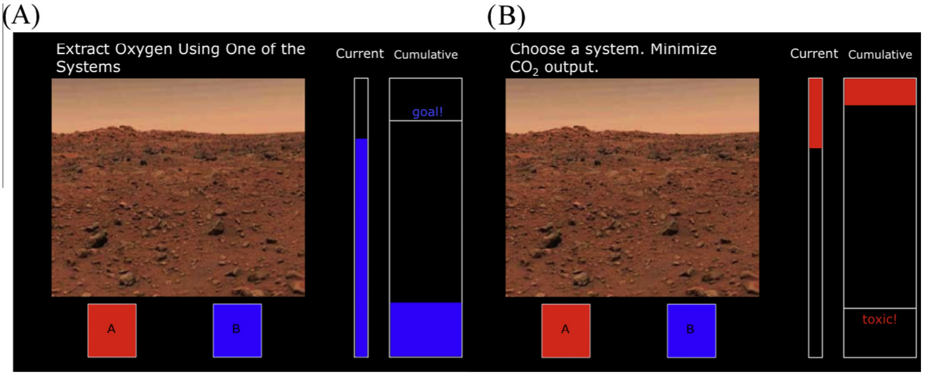

Fig. 1. (A and B): Screen shots from the experiment. (A) (left) Shows a screen shot from the gain-maximization condition. (B) (right) Shows a screen shot from the loss-minimization condition.

emit carbon dioxide at different rates. Participants were told only that their goal was to minimize the amount of carbon dioxide emitted in the task. A bar, representing a small carbon dioxide tank, displayed the amount of carbon dioxide emitted on each trial. This was then transferred into the larger tank and the next trial began. A line on the larger tank corresponded to the amount of carbon dioxide that could not be exceeded to sustain life on Mars, and participants were told to try to emit as little carbon dioxide as possible over the course of the experiment. The goal line again corresponded to selecting the optimal choice (increasing option) on approximately 80% of trials. Importantly, the gain-maximization and loss-minimization conditions are "yoked" in the sense that the losses in the loss-minimization condition are derived directly from the gains in the gain-maximization condition by subtracting a constant (100) from each reward value in Fig. 2A. Thus, the optimal task strategy is equated across gain-maximization and loss-minimization conditions.

A goal-directed strategy should lead to better performance in both the gain-maximization and loss-minimization versions of this task, as compared to a reward-based strategy, because participants using a goal-directed strategy should be more likely to select the increasing option to improve their state on future trials.

3. Results

To allow direct comparisons across the gain-maximization and loss-minimization versions of the task, we added 100 points to the outcome on each trial in the loss-minimization condition. We also examined the proportion of increasing (optimal) option selections. Performance for each group is displayed in Fig. 3.

We evaluated performance using a 2 (Condition: gain vs. loss) × 2 (RFQ: chronic promotion vs. chronic prevention) ANOVA on performance using participants' total points (Fig. 3, left). The main effect of condition was significant, $F(1,76) = 8.468$, $p = .005$, partial-$\eta^2 = .100$, indicating that overall

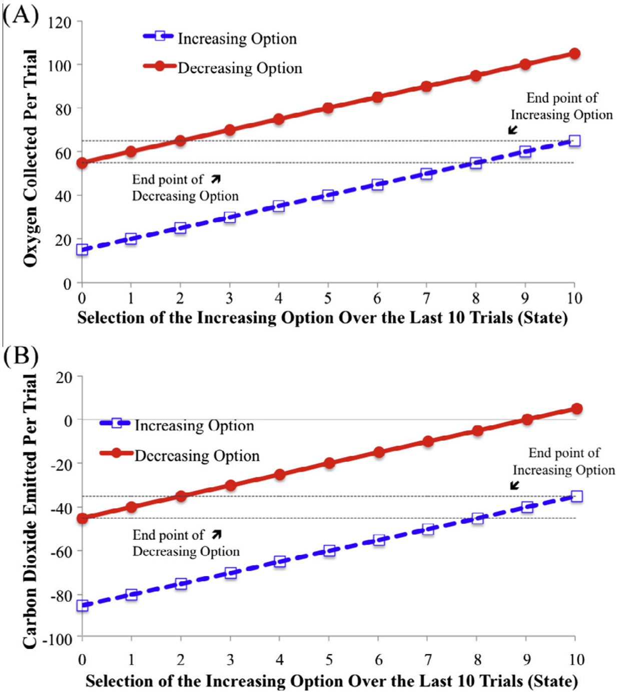

Fig. 2. (A and B): Experiment reward structures. (A) Shows the oxygen collected per trial in the gain-maximization task, which varied as a function of the number of times the Increasing Option was selected over the previous 10 trials. (B) Shows the carbon dioxide emitted per trial in the loss-minimization task, which also varied as a function of the number of times the Increasing Option was selected over the previous 10 trials.

participants performed better in the gain-maximization condition than the loss-minimization condition. The main effect of regulatory focus group was not significant, $F(1,76) = .007$, $p = .933$, partial-$\eta^2 < .001$, indicating that neither promotion nor prevention focused individuals held an overall advantage in the task. Importantly, we found a significant interaction, $F(1,76) = 14.106$, $p < .001$, partial-$\eta^2 = .157$ between regulatory focus (indicated by the RFQ) and task framing (gain-maximization or loss-minimization).

In decomposing this interaction, we found that in the gain-maximization condition, promotion-focused individuals (i.e., those in a regulatory fit; $M = 15{,}156$) performed better than prevention-focused individuals (i.e., those in a regulatory mismatch; $M = 14{,}745$), $t(38) = 2.876$, $p = .007$, Cohen's $d = .933$. In the loss-minimization condition, prevention-focused individuals (i.e. those in a regulatory fit; $M = 14{,}840$) performed better than promotion-focused individuals (i.e. those in a regulatory mismatch; $M = 14{,}409$), $t(38) = 2.494$, $p = .017$, Cohen's $d = .809$. We also compared performance within each regulatory focus group. Promotion-focused individuals performed better in gain-maximization than in loss-minimization, $t(38) = 4.989$, $p < .001$, Cohen's $d = 1.619$. Prevention-focused individuals' performance was numerically higher in loss-minimization than in gain-maximization, although this difference was not significant, $p = .57$. Thus, we found that performance in both gain-maximization and loss-minimization conditions was affected by the chronic regulatory focus of the participants, with better performance from participants in a match between chronic regulatory focus and reward structure.

As points in this task were correlated with the proportion of trials on which the optimal choice was selected, $r(79) = .984$, $p < .001$, our findings for points were replicated in the proportion of optimal selections. To evaluate whether differences in performance emerged early or late in the task we evaluated the proportion of optimal selections in each 50-trial block (Fig. 3, right). In a 2 (RFQ) $\times$ 2

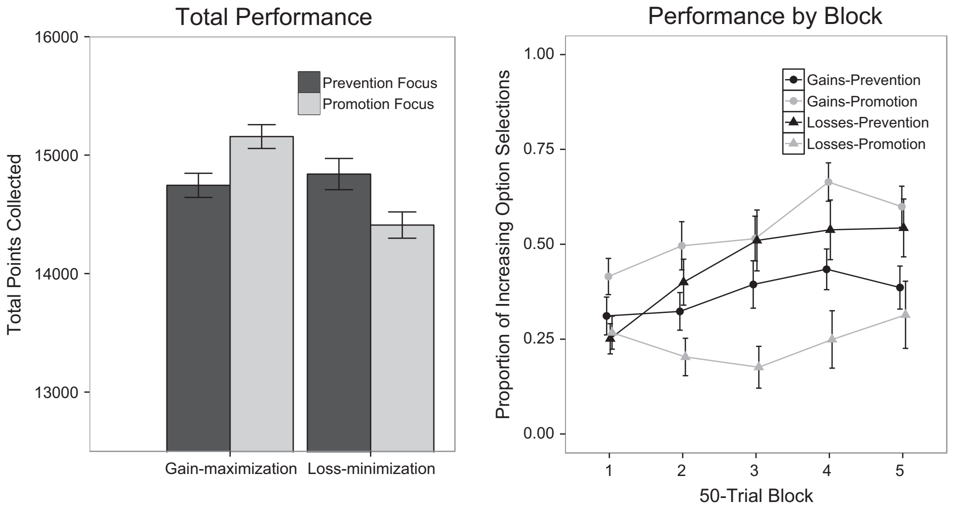

Fig. 3. Left: Task performance in all conditions. For the gain-maximization task, this was the total oxygen collected. For the loss-minimization task 100 points were added to the reward received on each trial to equate the yoked reward structures. Right: Performance across 50-trial block, measured using the proportion of trials on which the optimal (increasing) choice was selected. Error bars represent standard errors of the means.

(Condition) × 5 (Block) repeated measures ANOVA, we found a significant effect of block, $F(4,304) = 11.151$, $p < .001$, partial-$\eta^2 = .128$, indicating that participants' performance improved across the course of the experiment. We also found a significant Block × RFQ × Condition 3-way interaction, $F(4,304) = 4.264$, $p = .002$, partial-$\eta^2 = .053$, indicating that performance changed at different rates between our conditions. In the first block of 50 trials, condition and RFQ did not interact ($p = .335$). The interaction between condition and RFQ emerged in block 2 (trials 50–100) and persisted through block 5 ($ps \leqslant .002$). Very few participants (2.5%) reached the goal of selecting the optimal choice on 80% of all 250 trials, however, 43% of participants in a regulatory match and 18% of those in a mismatch reached this goal in at least one 50-trial block.

We next used participants' relative promotion focus scores (promotion focus score – prevention focus score) and conducted a linear regression to examine whether the association between task performance and regulatory focus score depended on the framing of the task. After mean-centering the regulatory focus difference scores and computing the RFQ-framing interaction term, the two predictors and the interaction were entered into a simultaneous regression model. This model explained a significant proportion of the variance in performance, $R^2 = .235$, $F(3,76) = 7.758$, $p < .001$. The interaction between RFQ score and task framing was significant, $\beta = .532$, $t(76) = 3.840$, $p < .001$, suggesting that the effect of regulatory focus score on performance depended on the framing of the task.

To interpret the interaction between RFQ score and task framing we examined simple correlations within each condition. In the gain-maximization condition, relative promotion focus scores (promotion focus – prevention focus) were positively correlated with performance, $r(39) = .446$, $p = .004$ (Fig. 4, left). Absolute promotion focus scores were also positively correlated with performance $r(39) = .413$, $p = .008$, while absolute prevention focus scores were negatively correlated with performance $r(39) = -.277$, $p = .083$. In the loss-minimization condition, relative promotion focus scores were negatively correlated with performance $r(39) = -.370$, $p = .019$ (Fig. 4, right). Although it did not reach significance, absolute promotion focus was negatively associated with performance $r(39) = -.211$, $p = .191$, and absolute prevention focus was positively associated with performance $r(39) = .235$, $p = .145$. Consistent with previous work (Higgins et al., 2001), prevention and promotion focus scores were not correlated $r(79) = -.033$, $p = .775$. While the prevention and promotion focus sub-scales were associated with performance, the strongest correlations were observed between the difference score that represents the relative promotion/prevention focus.

We also examined the effect of gender on performance to verify that different proportions of males and females in each group did not drive our effect. In a 2 (Condition) × 2 (RFQ) × 2 (Gender) ANOVA, we found that the main effect of gender was not significant ($p = .711$), the Gender × Condition was not significant ($p = .472$), the Gender × RFQ interaction was not significant ($p = .984$), and the Gender × Condition × RFQ interaction was not significant ($p = .735$), indicating that our observed group effects were not driven by the gender of our participants.

3.1. Computational modeling

One advantage of the task used in this experiment is that it is amenable to computational modeling that can extend our analysis beyond measures of performance to identify whether each group implemented similar or different strategies. In our study, differences in performance could be related to drastically different approaches (e.g. random responding) or slight shifts in strategy that can be captured by a single computational model. We fit the choice data for each participant on a trial-by-trial basis, assessing model fit by maximizing log-likelihood (see Eq. (10)). We compare the fits of three models: a baseline model, a basic reinforcement learning (RL) model, and a Hybrid Reinforcement-Learning model (Worthy et al., 2014). The baseline, or null, model is a simple one-parameter model that assumes that participants select each option with a constant probability. This model differs from the reinforcement learning models by assuming that choices on each trial are independent (i.e. participants do not learn across trials). Despite assuming that no learning takes place, this model can provide a good fit for participants in this task and is often used as a comparison model to evaluate the fit of learning models (Busemeyer & Stout, 2002; Gureckis & Love, 2009b; Worthy, Otto, & Maddox, 2012).

The basic RL model updates expected values for the chosen option on each trial based on the reward prediction error, and calculates probabilities for selecting each option with a Softmax decision rule (Sutton & Barto, 1998). The basic RL model only has two free parameters, a learning rate, and an inverse temperature (exploitation) parameter. This model can provide a good fit to data from this task, particularly when participants perform poorly (e.g. Worthy et al., 2012).

The Hybrid Reinforcement-Learning model is identical to the computational model implemented by Worthy et al. (2014) to assess the degree to which participants utilize model-free (habitual) versus model-based (goal-directed) decision-making in older and younger adults and similar to that recently used by Eppinger, Walter, Heekeren, and Li (2013). This model has also been shown to be a good fit for younger-adult decision-making data in our task (Worthy et al., 2014). The model includes the assumption that participants observe the hidden state on each trial (s_t), which corresponds to the number of times the increasing option has been selected over the previous 10 trials. The model values options based on two factors—(1) the probability of reaching a given state on the next trial (s_{t+1}) by selecting action a (the goal-directed component), and (2) the rewards that are experienced in each state (the habitual component). Several other models have assumed that subjects use similar information to determine behavior (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Eppinger et al., 2013; Gläscher et al., 2010; Gureckis & Love, 2009a). Following each trial, t, in state s_{t-1}, and arriving in state s, after taking action a, the model computes a state prediction error (SPE) using Eq. (1).

$$\delta_{\mathrm{SPE}} = 1 - T(s_{t-1}, a, s_t)_t \tag{1}$$

The model updates the state transition probability:

$$T(s_{t-1}, a, s_t)_{t+1} = T(s_{t-1}, a, s_t)_t + \eta\delta_{\mathrm{SPE}} \tag{2}$$

where η is a free parameter that represents the learning rate for state transition probabilities. State transition probabilities for all other states not arrived at on trial t, (s_t^*) are reduced according to Eq. (3), ensuring that all transition probabilities at a given state sum to 1. For example, if an action a was made in state 4 (corresponding to 4 increasing option selections in last 10 trials), arriving in

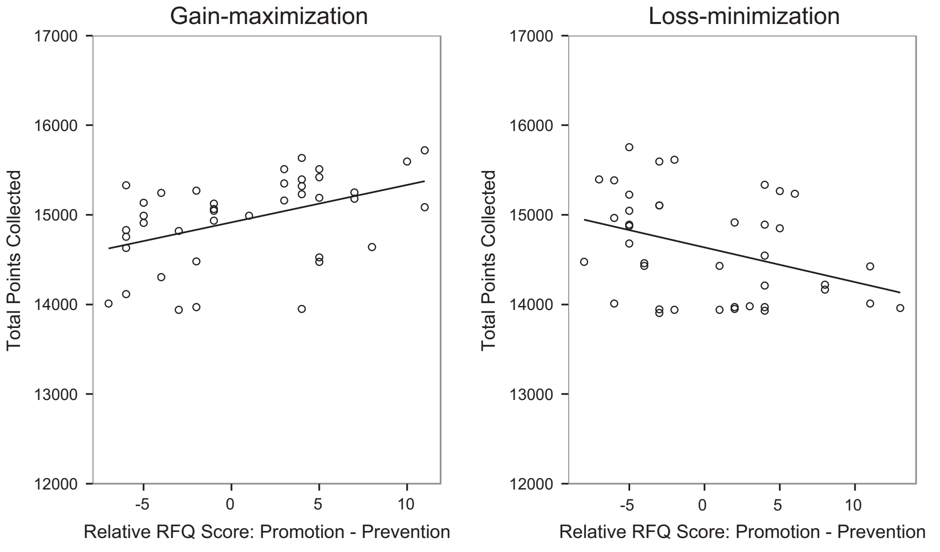

Fig. 4. Total points collected by relative regulatory focus (RFQ) score. The relative RFQ score is the difference between the promotion focus score and the prevention focus score. Positive values indicate higher scores in promotion focus than prevention focus, while negative values indicate higher scores in prevention focus than promotion focus.

state 5, the state transition probability of arriving at each other state from state 4, given action a, would be updated.

$$T(s_{t-1}, a, s_t^*)_{t+1} = T(s_{t-1}, a, s_t^*)_t \cdot (1 - \eta) \tag{3}$$

The model computes a state-based value for each action in each state (Model based value; $Q_{MB}(s, a)$) using a FORWARD learner that incorporates the state transition probabilities and the estimated maximum reward value for either option in the next state to determine the future value of each action (Eppinger et al., 2013; Gläscher et al., 2010).[2] In this task, there are three possible states that participants could transition to on the next trial (s_{t+1}): staying in the same state (e.g. participant is in state 4 and stays in state 4), moving up one state (e.g. moving to state 5 from state 4), or moving down one state (e.g. moving from state 4 to state 3). The Q_{MB} value for each state–action pair is estimated by the following equation:

$$Q_{MB}(s_t, a) = \sum_{(s-1)}^{(s+1)} T(s_t, a, s_{t+1}) * \max\,[Q_{MB}(s_{t+1}, a')] \tag{4}$$

Eq. (4) multiplies the probability of transitioning to each possible state on the next trial, having taken action a on trial t, by the maximum expected reward in state s_{t+1} for either action, a'.

The model tracks the reward-based expected reward values for each action in each state (Model free value; $Q_{MF}(s, a)$) using a state–action–reward–state–action (SARSA) learner (Gläscher et al., 2010; Morris, Nevet, Arkadir, Vaadia, & Bergman, 2006). The model computes the reward prediction error (δ_{RPE}) between the received reward, r, and the expected value on each trial (Eq. (5)) and updates the expected value for the current state–action pair, Eq. (6):

$$\delta_{RPE} = r - Q_{MF}(s_t, a)_t \tag{5}$$

[2] Prior work has defined future values over infinite horizons recursively using the Bellman equation (e.g. Gläscher et al., 2010). As noted in Eq. (4) below here we simply use the maximum model-free expected value from Eq. (6) below for selecting either action in the next state to estimate the model-based value of each action.

$$Q_{MF}(s_t, a)_{t+1} = Q_{MF}(s_t, a)_t + \alpha\delta_{RPE} \tag{6}$$

Eq. (6) utilizes another free parameter, α, which represents the learning rate for state–action pairs on each trial. For each state other than the state on the current trial (s_t^*), the Q_{MF} value for the selected action is updated in Eq. (7), where the degree to which the rewards received on each trial, r, are generalized to the same action in different states is represented by θ.

$$Q_{MF}(s_t^*, a)_{t+1} = Q_{MF}(s_t^*, a)_t + \theta(5 * r - Q_{MF}(s_t^*, a)) \tag{7}$$

The model then determines a net value for each action ($Q_{Net}(s,a)$) by taking a weighted average of the reward-based and state-based expected values in Eq. (10), where ω is a free parameter that determines the degree to which choices are based on the habitual versus goal-directed components of the model.

$$Q_{Net}(s_t, a) = \omega \cdot Q_{MB}(s_t, a) + (1 - \omega) \cdot Q_{MF}(s_t, a) \tag{8}$$

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

$$P(a, t) = \frac{e^{[\beta \cdot [Q_{Net}(s,a) + \pi \cdot rep(a)]]}}{\sum_{j=1}^{2} e^{[\beta \cdot [Q_{Net}(s,j) + \pi \cdot rep(j)]]}} \tag{9}$$

Here, β is an inverse temperature or exploitation parameter that determines the degree to which participants select the option with the highest expected value. Larger estimates of β indicate that the highest-valued option is more consistently selected, and values of β approaching 0 indicate that each option is being randomly selected. The autocorrelation parameter, π, accounts for tendencies to perseverate. Positive values of π indicate perseveration and negative values indicate switching regardless of trial outcome. 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 et al., 2013). In total, this model included six free parameters: η, α, θ, ω, β, and π.

Each of the three models was fit to each participant's data on a trial-by-trial basis, and the maximum likelihood was calculated for each model for each participant. For each model, we sought to estimate the parameters that maximized the likelihood of each participant's choices:

$$L_{model} = \prod_t P_{c,t} \tag{10}$$

where $P_{c,t}$ represents the probability of the model choosing the choice (c) made on trial t, given the participant's history of choices and rewards. The fit of these models was compared using Akaike Information Criterion (AIC; Akaike, 1974), which penalizes models based on their number of free parameters using the equation:

$$AIC_i = 2 \cdot V_i - 2 \cdot \ln(L_i) \tag{11}$$

where L_i is the maximum likelihood for model i and V_i is the number of free parameters in the model. Consistent with previous work (Gureckis & Love, 2009b; Worthy et al., 2012), we quantified the improvement in model fit provided by each learning model over the baseline model for each participant, corrected for differing numbers of free parameters, by subtracting the AIC of each learning model m from the AIC of the baseline model b. Table 2 shows the AIC^{b-m} score averaged across participants for each model in each experimental condition. Positive values of AIC^{b-m} indicate that the learning model provided a better fit than the baseline model. Table 2 displays in parenthesis the proportion of participants who were better fit by each learning model than the baseline model (positive values of AIC^{b-m}).

Table 2 also provides the difference in model fit using Bayesian Information Criterion (BIC; Schwarz, 1978) values as a goodness-of-fit metric. Similar to the AIC, the BIC is a model comparison metric based on the likelihood function of each model, but the BIC value calculation incorporates the number of observation, calculated using the equation:

$$BIC_i = k \cdot \ln(n) - 2 \cdot \ln(L_i) \tag{12}$$

where k is the number of free parameters in the model, and n is the number of trials being fit (250 in all cases). As with the AIC, we compared the BIC values of each learning model to the BIC of the baseline model, BIC^{b-m}.

3.1.1. Modeling results

For all conditions, the hybrid reinforcement-learning model provided a greater improvement over the baseline model than the basic RL model (Table 2). The hybrid RL model provided a better fit than the baseline model for a majority (97.5%) of participants. The BIC^{b-m} difference metric was positive in all conditions for the hybrid model. Using this comparison metric, the hybrid model again provided a better fit than the baseline model for the majority of participants (77.5%).

The comparison between the baseline model and learning models indicates that the majority of participants in each condition were best fit by the hybrid reinforcement-learning model, rather than being best fit by different models. To further examine the underlying causes of the observed performance differences, we examined the best-fitting values for the parameters of the hybrid model (Table 3) using a 2 (Condition) × 2 (RFQ) ANOVA. Three of the parameters (state learning rate η, reward learning rate α, and reward generalization rate θ) did not account for our observed performance differences. For these parameters RFQ × Condition interactions and main effects of RFQ and condition were all non-significant ($ps > .1$). The temperature parameter, β, was higher in the losses condition than in the gains condition, $F(1,76)$, = 12.331, $p = .001$, partial-$\eta^2 = .140$, which suggests that participants in the losses condition were more likely to choose the highest valued option. The RFQ × Reward interaction for the temperature parameter was not significant ($p = .462$). The perseveration parameter, π, was slightly higher in the gains condition than the losses condition, $F(1,76)$, = 2.892, $p = .093$, partial-$\eta^2 = .037$, which suggests that people were likely to switch more often when performing the task with losses compared to gains. We also observed a marginally significant RFQ × Condition interaction for this parameter $F(1,76)$, = 2.879, $p = .094$, partial-$\eta^2 = .036$. Promotion-focused individuals in the gains condition perseverated more than those in the losses condition ($p = .009$), while prevention focused individuals did not differ between conditions ($p = .998$).

We also examined the best-fitting values for the ω parameter, which estimated the weight given to the goal-directed component of the model. We found a significant Condition × RFQ interaction, $F(1,76) = 8.102$, $p = .006$, partial-$\eta^2 = .096$. The main effect of condition and RFQ were not significant ($ps > .69$). The ω parameter was highest in the conditions in which the reward structure and regulatory focus were in a match (promotion-focus in the gain-maximization condition, $M = .90$ and prevention-focus in the loss-minimization condition, $M = .86$) and lower in conditions in which reward structure and regulatory focus were in a mismatch (promotion focus in the loss-minimization condition, $M = .71$, and prevention-focus in the gain-maximization condition, $M = .72$). This difference was significant for the prevention-focused groups ($p = .023$) and marginally significant for the promotion-focused groups ($p = .09$). We examined the correlation between estimated ω parameter values performance over the course of the task. There was a strong positive association between task performance and the ω parameter, $r(79) = .602$, $p < .001$. This trend was consistent across all four conditions, $ps < .065$. This finding is consistent with Worthy et al. (2014), which found that the degree to which participants engaged in a goal-directed strategy, as indicated by the ω parameter, is positively correlated with performance in this task.

4. Discussion

This research builds on a regulatory fit framework proposed by Higgins (2000) and Maddox and Markman (2010) that emphasizes the notion that motivational influences on cognition emerge at multiple levels and are highly interactive. The effects of motivation at one level on cognitive processing can change substantially depending upon the motivational factors at another level. Understanding the interaction between chronic regulatory-focus mode and task framing (gain-maximization versus loss-minimization) is particularly important because it could provide a simple mechanism for optimizing performance in decision-making tasks by catering the framing to an individual's chronic motivational mode.

Table 2
Comparison of mean AIC^{b-m} and BIC^{b-m} score for each model, *m*, relative to the baseline model *b*. In parentheses is the percentage of participants for whom model *m* provides a better fit than baseline (i.e., AIC^{b-m} or BIC^{b-m} is positive). The best-fit model in each condition is indicated in bold.

	Gain-maximization		Loss-minimization	
	Promotion	Prevention	Promotion	Prevention
$AIC^{baseline}–AIC^{model}$				
Basic RL	−5.2 (.4)	8.7 (.5)	14.4 (.75)	10.3 (.6)
Hybrid	**100.8 (1)**	**71.7 (.95)**	**53.7 (1)**	**90.5 (.95)**
$BIC^{baseline}–BIC^{model}$				
Basic RL	−8.7 (.35)	5.2 (.45)	10.9 (.7)	6.8 (.45)
Hybrid	**83.2 (.8)**	**54.1 (.85)**	**35.1 (.65)**	**72.9 (.8)**

Table 3
Average best-fitting parameter estimates for each group.

	Gain-maximization				Loss-minimization			
	Promotion		Prevention		Promotion		Prevention	
State learning rate (η)	.42	(.37)	.51	(.41)	.44	(.46)	.33	(.39)
Reward learning rate (α)	.30	(.36)	.45	(.42)	.43	(.46)	.51	(.45)
Reward generalization rate (θ)	.12	(.28)	.33	(.44)	.19	(.27)	.17	(.32)
Model-based weight (ω)	.86	(.28)	.72	(.33)	.71	(.30)	.90	(.11)
Inverse temperature (β)	1.06	(1.68)	.90	(1.77)	2.96	(2.32)	2.14	(2.13)
Perseveration (π)	6.93	(6.48)	4.98	(12.01)	.27	(8.77)	4.97	(6.66)

Note: Standard deviations are listed in parenthesis.

We find that chronic regulatory focus interacts with task reward structure to induce regulatory fit. In a regulatory match (promotion-focused individuals in a gain-maximization task, or prevention-focused individuals in a loss-minimization task), participants perform better than those in a mismatch (promotion-focused individuals in a loss-minimization task, or prevention-focused individuals in a gain-maximization task). These results are consistent with previous work using a situational/induced regulatory focus in category learning (e.g. Grimm et al., 2008) and decision-making (Otto et al., 2010; Worthy et al., 2007). The present study also supports previous work suggesting that when the reward structure of the environment matches an individual's expectations, he or she utilizes full cognitive resources, but when the reward structure does not match an individual's expectations, he or she is likely to utilize fast-acting cognitive strategies that favor the most-salient alternatives (Markman et al., 2007; Worthy et al., 2007). In our case, a regulatory match resulted in increased use of goal-directed strategies in which individuals were motivated toward decisions that increased their future state. These computational modeling results suggest a mechanism for the behavioral effect that we observed and for those observed in previous studies.

These results could also be interpreted in the context of long-term versus short-term reward. Similar tasks have been used to assess preferences for long-term (or delayed) and short-term (or immediate) rewards (e.g. Gureckis & Love, 2009a; Otto & Love, 2010). Individuals in a regulatory match chose the better long-term option more frequently than those in a regulatory mismatch. Focusing on delayed reward relative to immediate reward takes additional cognitive resources (Bobova, Finn, Rickert, & Lucas, 2009; Worthy et al., 2012). Individuals in a regulatory match may have more cognitive resources available to use in the task, resulting in the ability to focus more on delayed rewards and better performance. Future work should examine whether these results extend to other tasks, such as the rising optimum task, that have also been widely used to evaluate preferences for immediate or delayed rewards (e.g. Bogacz, McClure, Li, Cohen, & Montague, 2007; Egelman, Person, & Montague, 1998; Montague & Berns, 2002). Similarly, future work should also examine the extent to which a regulatory match affects performance at different levels of task difficulty. Several variations of this decision-making task modify the level of difficulty, such as different separation between the reward options (Worthy et al., 2012), additional number of options (Worthy et al., 2014), noise (Gureckis & Love, 2009b), or specific cues about the current state (Gureckis & Love, 2009a), which could all provide future avenues of exploration.

One interesting aspect of the results of our study is that the effect of framing (gain-maximization or loss-minimization) had a stronger effect on promotion-focused individuals than prevention-focused individuals. That is, while prevention-oriented participants were better fit by parameters indicating more weight on state-based strategies, they performed relatively consistently across both gain-maximization and loss-minimization conditions, while promotion-focused individuals performed significantly better under gain-maximization than loss-minimization. This could suggest that promotion-oriented individuals may be more heavily influenced by reward framing in other contexts, or may be more affected by framing on a broader level (e.g. classic framing effects like those found in Tversky & Kahneman, 1981). One possible explanation for this reduced effect of framing on prevention-focused individuals is that payment for participation may have induced a temporary situational promotion focus that masked chronic focus effects. While other work (e.g. Higgins et al., 2001; Worthy et al., 2007) has found regulatory focus effects despite participant compensation, this possibility should still be examined in future work.

Importantly, we find no main effect of chronic motivational mode on performance, indicating that neither chronic promotion-focus nor chronic prevention-focus has an overall benefit (or disadvantage) for decision-making, as long as task framing is taken into account. However, if only a gain-maximization task had been considered we may have come to a very different conclusion and determined that those with a prevention-focus underperform in dynamic decision-making tasks. Our results indicate that performance can be maximized in decision-making when task structure is aligned with measures of chronic motivational mode.

Acknowledgments

This research was supported by the National Institute on Drug Abuse under Award Number DA032457 to WTM and the Air Force Office of Scientific Research, National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a, to JAC. We thank the Maddox Lab RAs for their help with data collection.

References

Aaker, J. L., & Lee, A. Y. (2001). "I" seek pleasures and "we" avoid pains: The role of self-regulatory goals in information processing and persuasion. *Journal of Consumer Research, 28*(1), 33–49.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. http://dx.doi.org/10.1109/TAC.1974.1100705.

Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bobova, L., Finn, P. R., Rickert, M. E., & Lucas, J. (2009). Disinhibitory psychopathology and delay discounting in alcohol dependence: Personality and cognitive correlates. *Experimental and Clinical Psychopharmacology, 17*(1), 51–61.

Bogacz, R., McClure, S. M., Li, J., Cohen, J. D., & Montague, P. R. (2007). Short-term memory traces for action bias in human reinforcement learning. *Brain Research, 1153*, 111–121.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436. http://dx.doi.org/10.1163/156856897X00357.

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., et al (2014). Mechanisms of motivation–cognition interaction: Challenges and opportunities. *Cognitive, Affective, & Behavioral Neuroscience, 14*(2), 443–472.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment, 14*(3), 253.

Camacho, C. J., Higgins, E. T., & Luger, L. (2003). Moral value transfer from regulatory fit: What feels right is right and what feels wrong is wrong. *Journal of Personality and Social Psychology, 84*(3), 498.

Cesario, J., Grant, H., & Higgins, E. T. (2004). Regulatory fit and persuasion: Transfer from "feeling right.". *Journal of Personality and Social Psychology, 86*(3), 388.

Cooper, J. A., Worthy, D. A., Gorlick, M. A., & Maddox, W. T. (2013). Scaffolding across the lifespan in history-dependent decision making. *Psychology and Aging, 28*(2), 505–514.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron, 69*(6), 1204–1215.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8*, 1704–1711.

Doya, K., Samejima, K., Katagiri, K., & Kawato, M. (2002). Multiple model based reinforcement learning. *Neural Computation, 14*, 1347–1369.

Egelman, D. M., Person, C., & Montague, P. R. (1998). A computational role for dopamine delivery in human decision-making. *Journal of Cognitive Neuroscience, 10*(5), 623–630.
Eppinger, B., Walter, M., Heekeren, H. R., & Li, S. C. (2013). Of goals and habits: Age-related and individual differences in goal-directed decision-making. *Frontiers in Neuroscience, 7*.
Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). Retrospective revaluation in sequential decision making: A tale of two systems. *Journal of Experimental Psychology: General, 143*(1), 182.
Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron, 66*(4), 585–595.
Grimm, L. R., Markman, A. B., Maddox, W. T., & Baldwin, G. C. (2008). Differential effects of regulatory fit on category learning. *Journal of Experimental Social Psychology, 44*, 920–927.
Gureckis, T. M., & Love, B. C. (2009a). Learning in noise: Dynamic decision-making in a variable environment. *Journal of Mathematical Psychology, 53*(3), 180–193.
Gureckis, T. M., & Love, B. C. (2009b). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition, 113*(3), 293–313.
Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1280–1300.
Higgins, E. T. (2000). Making a good decision: Value from fit. *American Psychologist, 55*, 1217–1230.
Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology, 31*, 3–23.
Maddox, W. T., Baldwin, G. C., & Markman, A. B. (2006). Regulatory focus effects on cognitive flexibility in rule-based classification learning. *Memory and Cognition, 34*(7), 1377–1397.
Maddox, W. T., & Markman, A. B. (2010). The motivation–cognition interface in learning and decision making. *Current Directions in Psychological Science, 19*(2), 106–110.
Markman, A. B., Maddox, W. T., Worthy, D. A., & Baldwin, G. C. (2007). Using regulatory focus to explore implicit and explicit processing. *Journal of Consciousness Studies, 14*, 132–155.
Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron, 36*(2), 265–284.
Morris, G., Nevet, A., Arkadir, D., Vaadia, E., & Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nature Neuroscience, 9*(8), 1057–1063.
Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science, 24*(5), 751–761.
Otto, A. R., & Love, B. C. (2010). You don't want to know what you're missing: When information about foregone rewards impedes dynamic decision making. *Judgment and Decision Making, 5*(1), 1–10.
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*(3), 797.
Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.
Roseman, I. J., Spindel, M. S., & Jose, P. E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology, 59*, 899–915.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.
Shah, J., Higgins, E. T., & Friedman, R. S. (1998). Performance incentives and means: How regulatory focus influences goal attainment. *Journal of Personality and Social Psychology, 74*, 285–293.
Sutton, R., & Barto, A. G. (1998). *Reinforcement learning*. Cambridge, MA: MIT Press.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.
Worthy, D. A., Cooper, J. A., Byrne, K. A., Gorlick, M. A., & Maddox, W. T. (2014). State-based versus reward-based motivation in younger and older adults. *Cognitive, Affective, & Behavioral Neuroscience, 14*(4), 1208–1220.
Worthy, D. A., Maddox, W. T., & Markman, A. B. (2007). Regulatory fit effects in a choice task. *Psychonomic Bulletin & Review, 14*(6), 1125–1132.
Worthy, D. A., Otto, A. R., & Maddox, W. T. (2012). Working-memory load and temporal myopia in dynamic decision making. *Journal of Experimental Psychology: Learning, Memory and Cognition, 38*(6), 1640–1658.