Regulatory fit effects in a choice task

DARRELL A. WORTHY, W. TODD MADDOX, AND ARTHUR B. MARKMAN
University of Texas, Austin, Texas

This article examines the interface between motivation and choice. In category learning, a regulatory fit has been shown to increase exploration of alternative response strategies even when exploration is suboptimal. In the present study, promotion- and prevention-focused subjects performed a choice task that required them to choose from one of two decks of cards on each trial. They either gained or lost points with each draw. In Experiment 1, optimal performance required an exploratory response pattern that entailed sampling from a deck that initially appeared disadvantageous but ultimately became advantageous. In Experiment 2, optimal performance required an exploitative response pattern. A softmax action selection model that includes an exploitation parameter was applied to each subject's data from both experiments and revealed greater exploration of alternative strategies for people with a regulatory fit. This response strategy was optimal in Experiment 1 and led to superior performance, but was suboptimal in Experiment 2 and led to inferior performance.

Motivation is central to action. The motivation literature makes a distinction between approach goals—positive states that one wants to achieve—and avoidance goals—negative states that one wants to avoid (see, e.g., Carver & Scheier, 1998). Higgins (1987) proposed *regulatory focus theory*, which argues for psychological states of readiness for potential gains or losses that tune the sensitivity of the motivational system. Previous research suggests that motivational states interact with the reward structure of the environment to affect people's *cognitive flexibility*, or their willingness to explore versus exploit different response strategies used to solve a classification task (Maddox, Baldwin, & Markman, 2006; Markman, Baldwin, & Maddox, 2005). For example, when optimal responding involves exploring the space of possible rules to find a low salience rule that leads to perfect performance, subjects attempting to earn an entry into a drawing for cash (which induces a *promotion* focus) perform best when they gain points upon responding, and subjects attempting to avoid losing an entry into a drawing (which induces a *prevention* focus) perform best when they lose points upon responding. Alternatively, when optimal responding involves exploiting an obvious rule that leads to good performance, subjects in a promotion focus perform best when they lose points upon responding, and subjects in a prevention focus perform best when they gain points upon responding. Thus, subjects in a *regulatory fit* will be more flexible, or more willing to explore alternative response strategies, whereas subjects in a *regulatory mismatch* will be less flexible, or less willing to explore alternative response strategies, regardless of whether greater flexibility leads to optimal performance. This hypothesis finds support in classification learning (Maddox et al., 2006; Markman et al., 2005).

In this article, we extend previous research in two ways. First, we explore these effects in choice behavior, and second, we implement a reinforcement-learning model that supports a more direct assessment of a given subject's exploration or exploitation of alternative choices in the environment.

In an uncertain environment, a decision maker is faced with an "exploration–exploitation" dilemma, in which he or she must balance the opposing demands of gathering and exploiting information (Aston-Jones & Cohen, 2005; Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Ishii, Yoshida, & Yoshimoto, 2002). Recent model-based analyses suggest that a parameter in the softmax action selection model captures the global propensity to explore (Aston-Jones & Cohen, 2005; Daw et al., 2006; Ishii et al., 2002). None of this work directly addresses how motivational factors might influence an agent's decision to explore or exploit alternatives in the environment. We hypothesized that the interaction between one's motivational state and the reward structure of the environment would influence the resolution of the exploration–exploitation dilemma, such that decision makers in a regulatory fit would more often choose to explore the environment, and decision makers in a regulatory mismatch would more often choose to exploit the option with the highest estimated value.

EXPERIMENT 1

We developed a choice task that was inspired by the Iowa gambling task (Bechara, A. R. Damasio, H. Damasio, & Anderson, 1994; Bechara, H. Damasio, & A. R. Damasio, 2000), in which subjects chose from one of two decks of cards, one that initially gave good values, but was ultimately worse to choose from (called the *disadvantageous*

W. T. Maddox, maddox@psy.utexas.edu

deck), and the other that initially gave bad values, but was ultimately better to choose from (called the *advantageous* deck; see Weber, Shafir, & Blais, 2004, for a similar task).

Subjects in the gains condition started with zero points and gained points on each draw, whereas subjects in the losses condition started with zero points and lost points on each draw. The decks were constructed so that a minimum of 25 cards had to be drawn from the advantageous deck to achieve the bonus criterion (see Figure 1). A subject who consistently exploited the disadvantageous deck, which initially appeared to be the better deck, would fail to achieve the bonus. To reach the bonus criterion, a subject had to be willing to sample from the advantageous deck even though the first cards drawn from it were of lower value than those drawn from the disadvantageous deck. This strategy required a willingness to explore different alternatives in the environment, and a willingness to avoid persistent use of the most salient response strategy.

Subjects in a *promotion* focus were told that they would receive an entry into a drawing for a 1 in 10 chance at winning $50 if they achieved the bonus criterion. Subjects in a *prevention* focus were given an entry into the drawing and told that they had to achieve the bonus criterion to avoid losing the entry.

We hypothesized that subjects in a regulatory fit (a promotion focus with a gains reward structure or a prevention focus with a losses reward structure) would explore more alternatives in the environment and thus be more willing to sample cards from the advantageous deck, which did not at first give better values. In contrast, we hypothesized that subjects in a regulatory mismatch (a promotion focus with a losses reward structure or a prevention focus with a gains reward structure) would repeatedly choose cards from the deck that initially gave the best values (the disadvantageous deck), and be less willing to sample from the deck that initially gave the worst values (the advantageous deck). This hypothesis followed from our work examining the effects of regulatory fit on classification learning (see Maddox et al., 2006).

Method

Forty undergraduates from the University of Texas community participated in the experiment for course credit or were paid $6 for their time. They were also given the opportunity to earn an entry into a drawing for $50 cash, and were told that no more than 10 participants would be included in each drawing. The two between-subjects independent variables were the situational regulatory focus (promotion and prevention) and the reward structure of the task (gains and losses).

The subjects were told that they would either earn (promotion focus) or keep (prevention focus) their ticket to the drawing if they met the bonus criterion. Two decks of cards appeared, one on the right and one on the left side of the screen. The subjects were asked to press a button on the left to draw from the deck on the left, and to press a button on the right to draw from a deck on the right. There were 80 cards in each deck, and each subject drew a total of 80 cards. On each trial, only the card that was drawn was discarded. In the gains condition, the disadvantageous deck gave an average value of 8 points per card over the first 30 cards drawn from it, an average value of 5 points over the next 20 cards drawn from it, and an average value of 2 points over the last 30 cards drawn from it, whereas the advantageous deck gave an average value of 3 points over the first 20 cards drawn from it, an average value of 7 points over the next 50 cards drawn from it, and an average value of 3 points over the last 10 cards drawn from it. The deck values were in the same order for each subject, and were equated for gains and losses by subtracting 11 from each gains condition deck value to produce the losses condition deck value (e.g., if a card gave a value of 7 in the gains condition, then the corresponding card in the losses condition gave a value of −4).

On each trial, subjects received between 1 and 10 points in the gains condition, and between −10 and −1 points in the losses condition. The bonus criterion was equated for gains and losses so that more than 25 cards had to be drawn from the advantageous deck to achieve the bonus. In the gains condition, subjects had to earn 450 points, and in the losses condition, subjects had to lose no more than 430 points to achieve the bonus criterion. The points required to meet the bonus criterion were displayed at the bottom of the screen, and the subject's current point total was displayed at the center of the screen.
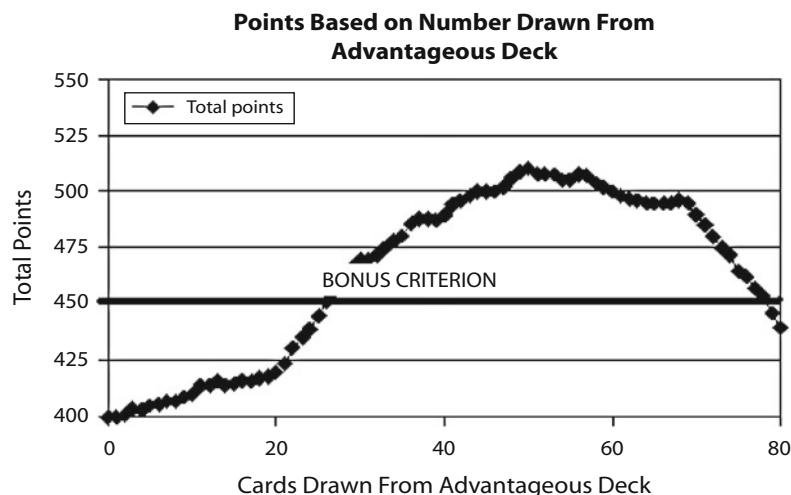

Figure 1. Total points earned with a gains payoff structure based on the number of cards drawn from the advantageous deck.

After 80 trials, subjects were given feedback on whether they had met the bonus criterion or not. If they met the bonus criterion, subjects in the situational promotion focus condition were given a ticket and told to enter it in the drawing, and subjects in the situational prevention focus condition were told that they could keep their ticket and enter it in the drawing.

Results

Performance measures. Figure 2A displays the average distance from the bonus criterion for each condition. A 2 (regulatory focus) × 2 (payoff structure) ANOVA revealed a significant interaction [$F(1,38) = 4.44$, $p < .05$, $\eta^2 = .11$]. As predicted, subjects in a regulatory fit were significantly closer to the criterion than were subjects in a regulatory mismatch. For the gains reward structure, subjects given a promotion focus ($M = -3.6$) were significantly closer to the bonus criterion than were subjects given a prevention focus ($M = -27.9$) [$F(1,18) = 6.05$, $p < .05$, $\eta^2 = .25$]. For the losses reward structure, subjects in a prevention focus ($M = -4.7$) performed better than did those with a promotion focus ($M = 14.6$), but this difference did not approach significance.

Subjects in a regulatory fit also drew more cards from the advantageous deck, which initially gave poorer values, but gave increasingly better values as more cards were drawn from the deck (Figure 2B). We conducted a 2 (regulatory focus) × 2 (reward structure) ANOVA on the number of cards drawn from the advantageous deck. The interaction was significant [$F(1,38) = 4.14$, $p < .05$, $\eta^2 = .10$]. For the gains reward structure, subjects given a promotion focus ($M = 26.3$) drew significantly more cards from the advantageous deck than did subjects given a prevention focus ($M = 20.1$) [$F(1,18) = 5.48$, $p < .05$, $\eta^2 = .23$]. In the losses reward structure, although subjects in a prevention focus ($M = 26.2$) drew more cards from the advantageous deck than did subjects in a promotion focus ($M = 22.2$), the effect was not significant. As predicted, subjects in a regulatory fit were more willing to sample from the advantageous deck than were subjects in a regulatory mismatch.

Model-based analyses. We implemented a version of the softmax action selection model in which the probability of selecting an alternative varies as a function of its
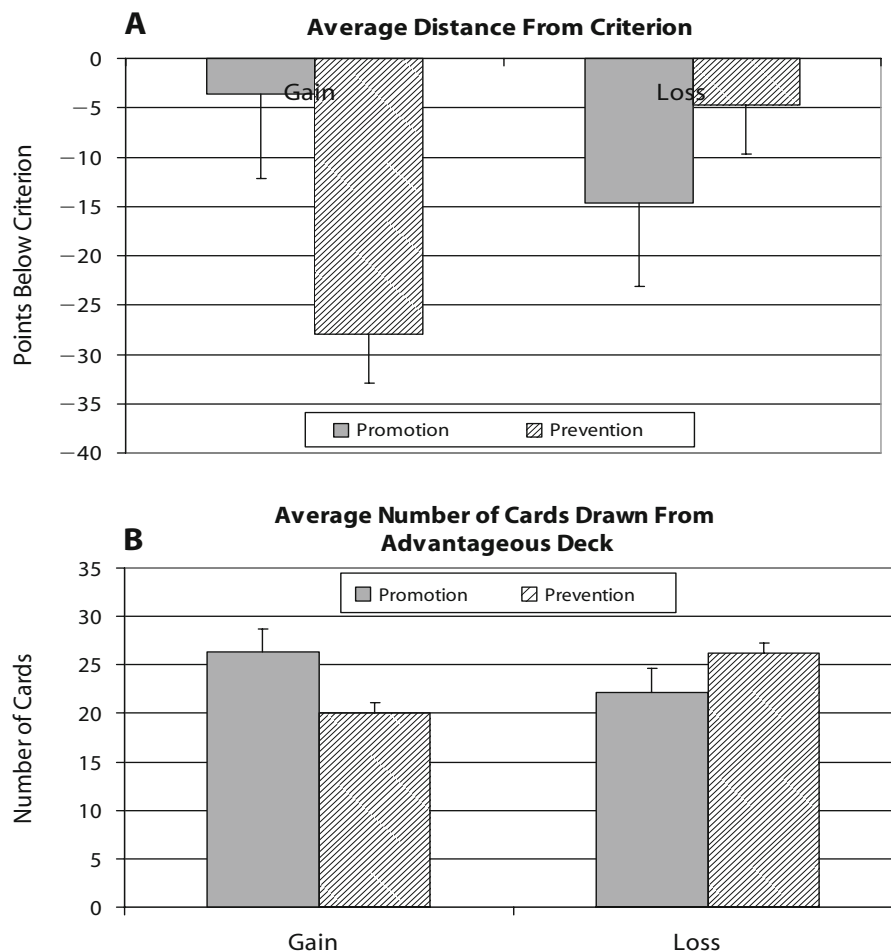

Figure 2. (A) Average distance from the bonus criterion for each group. Distances were equated for gain and losses payoff structures. (B) Average number of cards drawn from advantageous deck for each group.

estimated value (Daw et al., 2006; Sutton & Barto, 1998). In this model, the probability of choosing option a on trial t is:

$$P_{a,t} = \frac{e^{[\gamma E_t(a)]}}{\sum_{b=1}^n e^{[\gamma E_t(b)]}}, \quad (1)$$

where γ^1 is an exploitation parameter, and $E_t(a)$ is an estimate of the reward associated with choosing from deck a on trial t. In Equation 1, as $\gamma \to \infty$, the option with the higher estimated reward is exploited, whereas, as $\gamma \to 0$, response selection becomes more exploratory. We used an incremental update rule for updating an average E_k of the k past (r) rewards:

$$E_{k+1} = E_k + \alpha(r_{k+1} - E_k), \quad (2)$$

where α, a recency parameter, varies from 0 to 1. When $\alpha = 1$, Equation 2 reduces to

$$E_{k+1} = r_{k+1}, \quad (3)$$

so that only the most recent rewards are used to estimate the value of a response option, and as $\alpha \to 0$, Equation 2 reduces to

$$E_{k+1} = E_k, \quad (4)$$

so that all previous rewards are equally weighted.

We used the models described above to evaluate the degree to which subjects explored or exploited response alternatives in the environment. The model accounted for 79% of responses across all subjects. We hypothesized that subjects in a regulatory fit would be more exploratory and thus yield lower estimates of the exploitation parameter than were subjects in a regulatory mismatch. Figure 3A displays the average exploitation parameter values for each condition. A 2 (regulatory focus) × 2 (reward structure) ANOVA conducted on the exploitation parameter values revealed a significant interaction between regulatory focus and reward structure [$F(1,38) = 9.99$, $p < .01$, $\eta^2 = .22$]. For the gains reward structure, subjects in a promotion focus ($M = .34$) had significantly lower exploitation parameter values than did subjects in a prevention focus ($M = .53$) [$F(1,18) = 5.53$, $p < .05$, $\eta^2 = .24$]. For the losses reward structure, subjects in a prevention focus ($M = .27$) had significantly lower exploitation parameter values than did subjects in a promotion focus ($M = .51$) [$F(1,18) = 4.83$, $p < .05$, $\eta^2 = .21$].
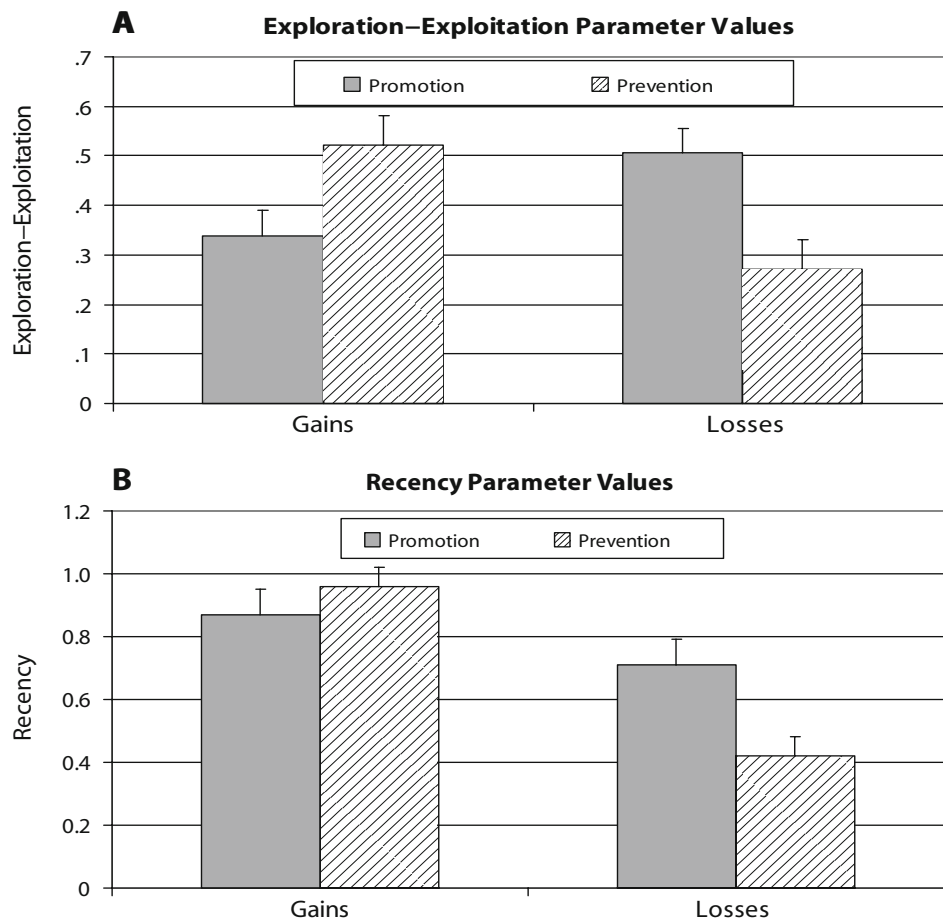

Figure 3. (A) Average best-fitting exploration–exploitation parameter values for each condition in Experiment 1. (B) Average best-fitting recency parameter values for each condition in Experiment 1.

We also examined the best-fitting recency parameter for each subject. Figure 3B shows the average best-fitting recency parameter values for each condition. A 2 (regulatory focus) $\times$ 2 (reward structure) ANOVA conducted on the recency parameter values revealed a significant main effect for reward structure [$F(1,38) = 12.88, p < .01, \eta^2 = .26$]. Subjects with a gains reward structure tended to give greater weight to recent gains, whereas subjects with a losses reward structure tended to give less weight to recent losses. There was a marginal interaction between regulatory focus and reward structure [$F(1,38) = 4.05, p = .052, \eta^2 = .10$]. Subjects in a regulatory fit, on average, had lower recency parameter values than did subjects in a regulatory mismatch. However, there was no significant difference in recency parameter values between promotion ($M = .87$) and prevention ($M = .96$) subjects in the gains condition, and the difference in recency parameter values between promotion ($M = .72$) and prevention (.43) subjects in the losses condition was only marginally significant [$F(1,38) = 2.90, p = .11, \eta^2 = .14$].[2]

Discussion

The results from Experiment 1 mirror those found in category learning. When exploring alternative strategies or response options is advantageous, as in the present choice task, subjects in a regulatory fit behaved more optimally than did subjects in a regulatory mismatch. In the present domain of choice, we were able to develop and test a model that included a specific exploitation parameter that provided a direct numerical index of how subjects were handling the exploration–exploitation dilemma. The results indicate that subjects in a regulatory fit had a stronger tendency to explore alternative response strategies, whereas subjects in a regulatory mismatch had a stronger tendency to exploit the alternative with the higher predicted value.

One might argue that subjects in a regulatory fit may be better equipped to solve a task than subjects in a regulatory mismatch, and not simply more prone to explore alternative response strategies. We cannot rule out this possibility in Experiment 1, because optimal performance required exploration. As a test of these two alternatives, we examined the effects of regulatory fit in a task similar to Experiment 1, but for which optimal performance required an exploitative strategy. If a regulatory fit leads to increased exploration, it should adversely affect performance. On the other hand, if a regulatory fit simply leads to better performance, it should improve performance.

EXPERIMENT 2

In Experiment 2, we used a choice task for which exploration of the deck with a lower expected value led to suboptimal performance.

Method

Thirty undergraduates from the University of Texas at Austin participated in the experiment for course credit or were paid $6 for their time. As in Experiment 1, subjects were told that they would either earn (promotion focus) or keep (prevention focus) their ticket to the drawing if they met the bonus criterion. Regulatory focus was the only independent variable, because all subjects had a gains reward structure, in which they gained points on each trial.

In Experiment 1, the cards from the unselected deck were not discarded, but in Experiment 2, one card from both the selected and unselected decks was discarded on every trial, although subjects were not given instructions about which cards were discarded. Deck A gave values that averaged 3 points during the first 30 trials, 4 points over the next 20 trials, and 7 points over the last 30 trials, whereas Deck B gave values that averaged 8 points over the first 30 trials, 6 points over the next 20 trials, and 3 points over the last 30 trials. Thus, the optimal strategy was to draw from Deck B during the first 50 trials and to draw from Deck A over the last 30 trials. A participant using such a strategy would earn 570 points. We set the bonus criterion at 550 points so that a failure to exploit the deck that currently gave the highest reward would lead to suboptimal performance in the task. With the exception of the deck characteristics and bonus criterion described above, all other procedures were the same as in Experiment 1.

Results

Performance measures. Figure 4A shows the distance from criterion for each condition in Experiment 2. As predicted from our hypothesis, subjects in a prevention focus with a gains reward structure (a regulatory mismatch) were significantly closer to the criterion ($M = -26.6$) than were subjects with a promotion focus ($M = -76.1$) [$F(1,28) = 8.3, p < .01, \eta^2 = .23$]. Figure 2B shows the average number of cards drawn from Deck A during each epoch of the task. A repeated measures ANOVA with epoch as a within-subjects factor and regulatory focus as a between-subjects factor revealed a main effect of regulatory focus [$F(1,28) = 9.42, p < .01, \eta^2 = .25$]. Pairwise comparisons were performed for each epoch. Subjects in a promotion ($M = 10.1$) focus drew significantly more cards from Deck A during the first 30 trials than did subjects in a prevention ($M = 4.1$) focus [$F(1,28) = 10.49, p < .01, \eta^2 = .27$]. During Trials 31–50, subjects in a promotion ($M = 6.5$) focus continued to draw more cards from Deck A than did subjects in a prevention ($M = 3.1$) focus [$F(1,28) = 5.7, p < .05, \eta^2 = .16$]. During the final epoch of trials, when Deck A became more advantageous to draw from than Deck B, subjects in a prevention ($M = 25.5$) focus drew more cards from Deck A than did subjects in a promotion ($M = 22.6$) focus, although the effect was not significant.

Model-based analyses. Fits of the softmax model described above also supported our hypothesis. Across both conditions, the model accounted for 85% of responses. As in Experiment 1, subjects in a regulatory fit ($M = .042$) had significantly lower exploitation parameter values than did subjects in a regulatory mismatch ($M = .80$) [$F(1,28) = 6.22, p < .05, \eta^2 = .18$]. Thus, the decision-making behavior for subjects in both a regulatory fit and a regulatory mismatch was the same for both tasks, even though exploration was optimal in Experiment 1 and exploitation was optimal in Experiment 2.

We also analyzed recency parameter values for subjects in a promotion and a prevention focus, but found no significant differences between promotion and prevention subjects.[3]

A   **Average Distance From Criterion**
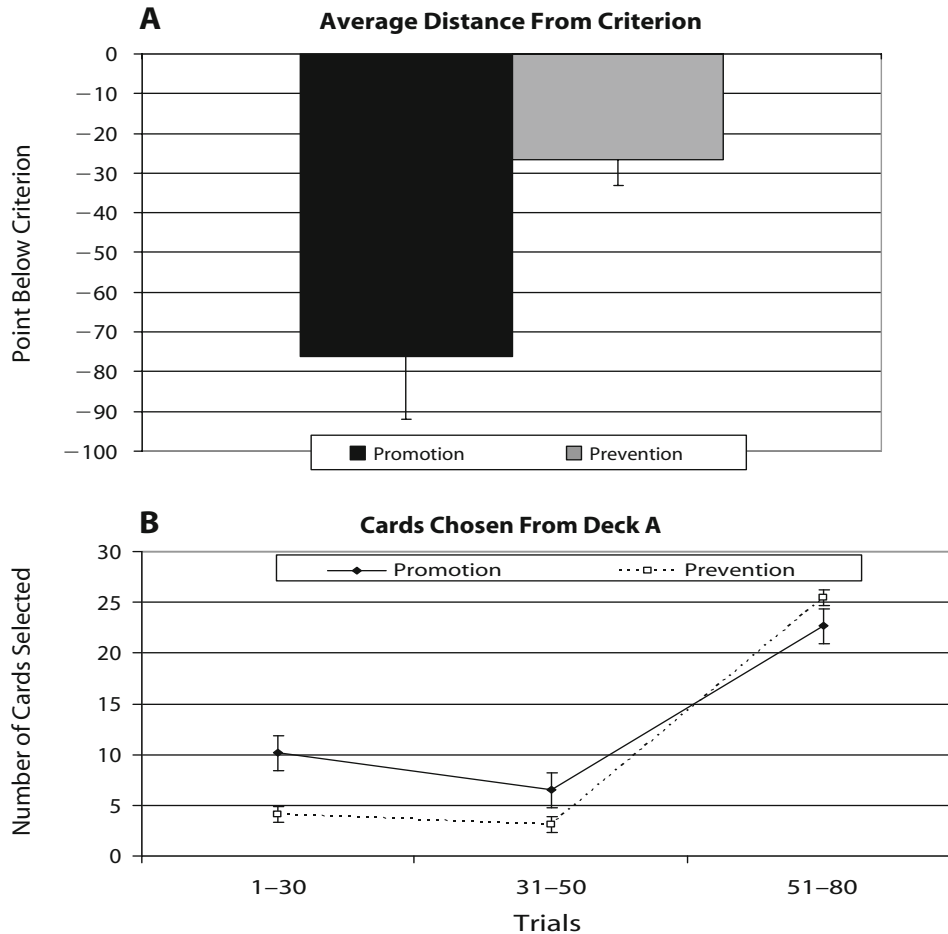

B   **Cards Chosen From Deck A**

Figure 4. (A) Average distance from criterion for promotion and prevention subjects in Experiment 2. (B) Average number of cards drawn for promotion and prevention subjects.

GENERAL DISCUSSION

This article examines the interface between motivation and choice. Of great interest in the reinforcement-learning and decision-making literatures is the exploration–exploitation dilemma (Aston-Jones & Cohen, 2005; Daw et al., 2006). Previous research has addressed the neural basis of exploration versus exploitation, but little work has examined causal factors that influence the resolution of this dilemma, such as motivational factors. Here, we have shown that the interaction between two factors in the environment, regulatory focus and reward structure, strongly influences the degree to which one explores or exploits alternative options in the environment. Subjects whose regulatory focus matched the reward structure of the task (a regulatory fit) were more prone to explore options with a lower estimated value, whereas subjects whose regulatory focus did not match the reward structure of the task were more likely to exploit the alternative with the highest estimated value. This occurred in situations in which exploration led to more optimal performance (Experiment 1), and in situations in which exploration led to less optimal performance (Experiment 2).

It is important to note that the resolution of the exploration–exploitation dilemma did not merely depend on whether subjects received gains or losses. Rather, the influence of gains and losses on choice depended completely on the situational regulatory focus. In Experiment 1, when there was a match between regulatory focus and reward structure, performance was better, and subjects were more willing to explore alternative options. Subjects who were in a regulatory fit were more willing to avoid persistent selection from the deck that was initially advantageous, and were willing to try alternative response strategies. In Experiment 2, subjects in a regulatory fit showed exploratory behavior similar to subjects in Experiment 1, even though this behavior led to less optimal performance.

One might ask *why* a match between one's regulatory focus and the reward structure of the environment leads to more exploratory behavior. Elsewhere (Markman, Maddox, Worthy, & Baldwin, 2007), we have suggested that when the reward structure of the environment matches an individual's expectations, he or she should bring his or her full cognitive resources to bear on problems to be solved in that environment. However, when the reward
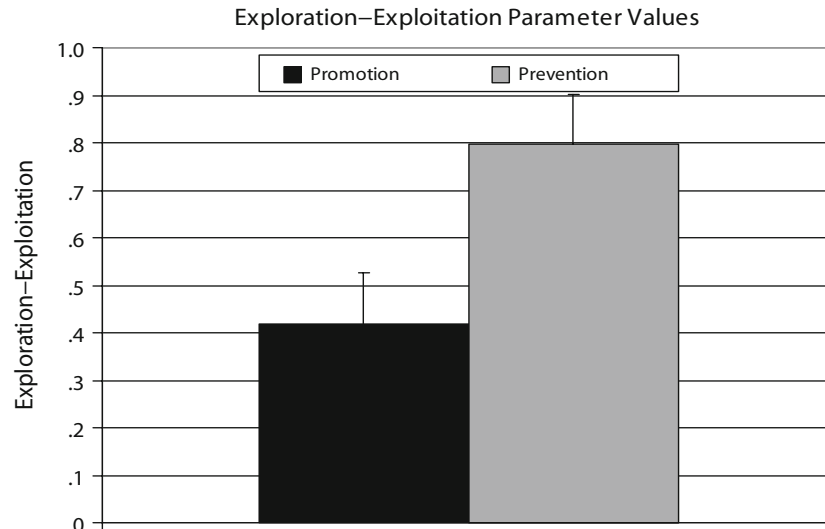
Exploration–Exploitation Parameter Values

Figure 5. Average best-fitting exploration–exploitation parameter values for each condition in Experiment 2.

structure does not match an individual's expectations, he or she is likely to engage fast-acting cognitive strategies until the environment can be better understood. Bringing one's full cognitive resources to bear on solving a problem will likely lead to the implementation of a greater variety of strategies, whereas engaging fast-acting strategies will likely lead to exploitation of the most salient alternatives. Although this is a reasonable hypothesis as to why a regulatory fit produces these effects, it is, nevertheless, only speculative at this point.

Regulatory focus may alter the behavioral patterns in tasks in other domains, as well. Recent work on foraging behavior found evidence of undermatching in groups of humans foraging between resource pools that differed in their output (Goldstone & Ashpole, 2004; Goldstone, Ashpole, & Roberts, 2005). However, motivational factors were not explicitly controlled. In fact, in these studies, subjects appeared to be in a regulatory fit because they were attempting to increase their chances of winning a drawing for cash (a promotion focus) while gaining points by foraging for food in the task (a gains rewards structure). Our present findings suggest that a regulatory fit leads to increased exploration of response alternatives, and they may partially explain the degree of undermatching found in these foraging experiments. Our findings also lead to the intriguing hypothesis that a regulatory mismatch might lead to an increase in exploitation of resource pools that give food more frequently in human foraging experiments.

We have shown that motivational factors in the environment influence the resolution of the exploration–exploitation dilemma in a choice task. It is becoming clear that motivation strongly affects cognition and behavior.

REFERENCES

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, **28**, 403-450.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, **50**, 7-15.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making, and the orbitofrontal cortex. *Cerebral Cortex*, **10**, 295-307.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, **441**, 876-879.

Goldstone, R. L., & Ashpole, B. C. (2004). Human foraging behavior in a virtual environment. *Psychonomic Bulletin & Review*, **11**, 508-514.

Goldstone, R. L., Ashpole, B. C., & Roberts, M. E. (2005). Knowledge of resources and competitors in human foraging. *Psychonomic Bulletin & Review*, **12**, 81-87.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, **94**, 319-340.

Ishii, S., Yoshida, W., & Yoshimoto, J. (2002). Control of exploitation–exploration meta-parameter in reinforcement learning. *Neural Networks*, **15**, 665-687.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, **53**, 49-70.

Maddox, W. T., Baldwin, G. C., & Markman, A. B. (2006). A test of the regulatory fit hypothesis in perceptual classification learning. *Memory & Cognition*, **34**, 1377-1397.

Markman, A. B., Baldwin, G. C., & Maddox, W. T. (2005). The interaction of payoff structure and motivation in classification. *Psychological Science*, **16**, 852-855.

Markman, A. B., Maddox, W. T., Worthy, D. A., & Baldwin, G. C. (2007). Using regulatory focus to explore implicit and explicit processing in concept learning. *Journal of Consciousness Studies*, **14**, (9-10), 132-155.

SUTTON, R. S., & BARTO, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

WEBER, E. U., SHAFIR, S., & BLAIS, A. R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, **111**, 430-445.

YECHIAM, E., BUSEMEYER, J. R., STOUT, J. C., & BECHARA, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science*, **16**, 973-978.

NOTES

1. For a similar use of the γ parameter in a category-learning model, see Maddox and Ashby (1993). See also Yechiam, Busemeyer, Stout, and Bechara (2005) for a model similar to the softmax action selection model implemented here.

2. In Experiment 2, we find the opposite pattern, where subjects in a regulatory fit have higher recency parameter values than do subjects in a regulatory mismatch. We therefore conclude that the interaction between regulatory focus and reward structure does not have a consistent effect on this parameter.

3. In Experiment 1, we found that subjects in a regulatory fit had lower recency parameter values. However, in Experiment 2, subjects in a regulatory fit ($M = .83$), on average, had higher recency parameter values than did subjects in a regulatory mismatch ($M = .71$).