



Motivation Influences Choice in the Exploration/ Exploitation Dilemma: Regulatory Fit Effects in a Gambling Task



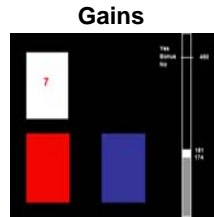
Darrell A. Worthy, W. Todd Maddox & Arthur B. Markman
University of Texas, Austin

Introduction

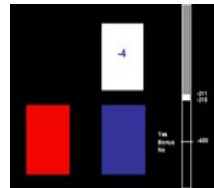
- In choice tasks one must decide whether to exploit the "best" option or explore other, less appealing options.
- In category learning, a *regulatory fit* has been shown to increase exploration of alternative response strategies even when exploration is sub-optimal.¹
- A regulatory fit occurs when one's regulatory focus (i.e. long-term goal) matches the reward structure of the task (i.e. short term goal).

Exploratory Task²

On each trial subjects drew a card



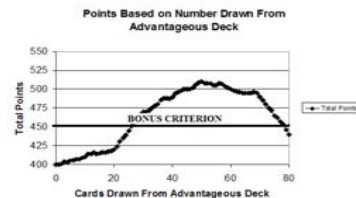
Gains



One deck, the 'advantageous deck', initially gave poor values, but gave better values as more cards were drawn from it

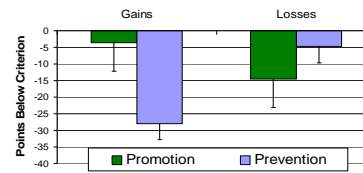
Advantageous	Disadvantageous
3 points over the first 20 cards drawn	8 points over the first 30 cards drawn
7 points over the next 50 cards drawn	5 points over the next 20 cards drawn
3 points over the last 10 cards drawn	2 points over the last 30 cards drawn

A minimum of 25 cards had to be drawn from the advantageous deck to reach the bonus



Results

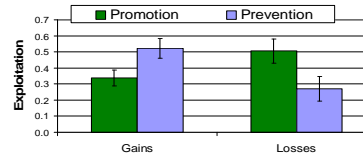
Average Distance from Criterion



Subjects in a regulatory fit were closer to the criterion than subjects in a regulatory mismatch

- We fit each subject's data using the Softmax action selection model.³
- This model estimates an exploitation parameter value that best fits the data.
- Higher values of this parameter indicate more exploitative responding.

Exploitation parameter



Subjects in a regulatory fit had lower exploitation parameter values than subjects in a mismatch

Exploitative Task

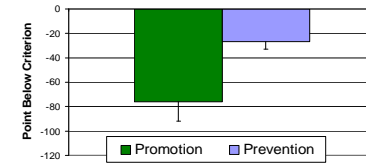
The optimal strategy required exploiting Deck B for the first 50 trials and Deck A for the last 30 trials.

Deck A	Deck B
3 points over the first 30 trials	8 points over the first 30 trials
4 points over the next 20 trials	6 points over the next 20 trials
7 points over the last 30 trials	3 points over the last 30 trials

All subjects had a gains reward matrix

Results

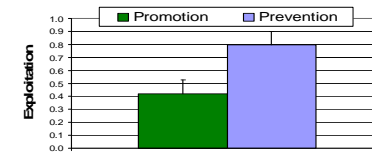
Average Distance from Criterion



Subjects in a regulatory fit were further from the criterion when the task favored an exploitative strategy.

We again fit the data using the Softmax model.

Exploitation Parameter



Subjects in a regulatory fit had lower exploitation parameter values than subjects in a mismatch

Discussion

- Motivational factors had a significant impact on the resolution of the Exploration/Exploitation Dilemma.
- Both behavioral and model-based analyses showed that a regulatory fit led to more exploratory behavior
- This effect occurred *regardless* of whether exploring or exploiting was more advantageous.
- It was the *interaction* between one's regulatory focus and the reward structure of the task which influenced behavior.

References

[1] 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.
 [2] Worthy, D.A., Maddox, W.T., & Markman, A.B. (in press). Regulatory Fit Effects in a choice task. *Psychonomic Bulletin and Review*.
 [3] Sutton, R.G., & Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. MIT Press. Cambridge, Massachusetts.
 Email: worthyda@mail.utexas.edu
<http://homepage.psy.utexas.edu/homepage/group/MaddoxLAB/>
 supported by AFOSR grant FA9550-06-1-0204, and NIMH grant MH077708 to WTM and ABM

Regulatory Focus

Promotion Prevention

	Promotion	Prevention
Gains	Fit	Mismatch
Losses	Mismatch	Fit

Hypotheses

Regulatory Focus

Promotion Prevention

	Promotion	Prevention
Gains	More Exploratory	More Exploitative
Losses	More Exploitative	More Exploratory

Softmax model

In this model the probability of choosing option a on trial t is:

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

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

$$E_{k+1} = E_k + \alpha[r_{k+1} - E_k]$$

where α , a recency parameter, varies from 0 to 1.

Note: The γ parameter is the exploitation parameter plotted on the poster. This was shown to be higher for subjects in a regulatory mismatch than subjects in a regulatory fit.