Modeling the Motivation-Learning Interface in Learning and Decision Making (FA9550-06-1-0204)

PI’s: W. Todd Maddox (University of Texas, Austin) & Arthur B. Markman (University of Texas, Austin)
## Motivation-Learning Interface (Maddox/Markman)

**Objective:**
To understand the influence of motivational incentives on learning and performance through empirical and computational model-based analyses.

To improve mathematical models of learning and performance based on data.

**Technical Approach:**
Manipulate people’s motivational state through global and local incentive manipulations.

Conduct experiments on choice and signal detection to understand how motivation affects the optimality of performance, and exploration/exploitation tradeoff.

**DoD Benefit:**
Motivational states guide actions but differ across military and non-military settings.

Goal is to identify motivational states that optimize performance in each setting.

Behavioral and model-based analyses illuminate these effects and characterize them along an exploration-exploitation continuum.

### Budget:

<table>
<thead>
<tr>
<th></th>
<th>FY06</th>
<th>FY07</th>
<th>FY08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual/Planned $K</td>
<td>$152</td>
<td>$152</td>
<td>$152</td>
</tr>
</tbody>
</table>

**Annual Progress Report Submitted?**  
Y   Y   N

**Project End Date:** 2/28/09
List of Project Goals

1. Develop and test a choice/gambling task

2. Examine and model motivational influences on this choice task

3. Examine and model variants of this task

4. Explore social influences on motivation, learning and performance

5. Extend to and model related tasks (signal detection, decision criterion learning, dynamic decision making)
Progress Towards Goals (or New Goals)

1. Develop and test a choice/gambling task  
   **Done**

2. Examine and model motivational influences on this choice task  
   **Initial studies completed and published**

3. Examine and model variants of this task  
   **Some competed and published; others in progress**

4. Explore social influences on motivation, learning, and performance  
   **Initial studies completed and published; others in progress**

5. Extend to and model related tasks (signal detection, decision criterion learning, dynamic decision making)  
   **Work in progress**
Research Questions

• What does it mean to “motivate” someone to do “well”? 
• How do we achieve this aim?
Layman’s Answer

• What does it mean to “motivate” someone to do “well”?
  – Get them to “try harder” (maximize number correct, targets destroyed, etc)

• How do we achieve this aim?
  – Give them an incentive for maximizing (raise, promotion, etc)

• Our research suggests that offering a global incentive (raise) for maximizing local incentive (number correct) is too simple a story and is misleading, even if we define “trying harder” as “attempting to respond optimally”.

Three-Factor Framework

- Influence of motivating incentives on performance involves a complex three-way interaction between three factors

  - **Global incentives** (Factor 1)
    - Approach some global reward (raise), or
    - Avoid losing some reward (avoid a pay cut)

  - **Local incentives** (Factor 2)
    - Maximize gains (maximize points earned)
    - Minimize losses (minimize points lost)

  - **Task demand** (What strategy is optimal?) (Factor 3) - Exploration or exploitation optimal
Overview of this talk

- Three factor (regulatory fit) framework
- Studies of choice
- Extensions of choice task and model
- Social influences on motivation
- Extensions to signal detection, decision criterion learning, dynamic decision making
## Global Incentives (Regulatory Focus)

<table>
<thead>
<tr>
<th>Approach (Promotion Focus)</th>
<th>Achieve Global Task Performance Criterion → Raffle ticket for $50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoidance (Prevention Focus)</td>
<td>Achieve Global Task Performance Criterion → Keep $50 raffle ticket given initially</td>
</tr>
</tbody>
</table>
## Task Reward Structure
(Local Trial-by-trial Task Goal)

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
</table>
| **Gains** | Earn points for all responses  
(Earn more points for correct choice than for incorrect choice) |
| **Losses** | Lose points for all responses  
(Lose fewer points for correct choice than for incorrect choice) |
Consider the bigger picture

<table>
<thead>
<tr>
<th>Promotion Focus</th>
<th>Gains</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit</td>
<td></td>
<td>Mismatch</td>
</tr>
<tr>
<td>Mismatch</td>
<td></td>
<td>Fit</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prevention Focus</th>
<th>Gains</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit</td>
<td></td>
<td>Mismatch</td>
</tr>
<tr>
<td>Mismatch</td>
<td></td>
<td>Fit</td>
</tr>
</tbody>
</table>

• Hypothesis: Fit increases exploration
• Exploration can be defined within tasks
  – Willingness to shift strategies
  – Willingness to explore a set of options
Consider the bigger picture

<table>
<thead>
<tr>
<th>Promotion Focus</th>
<th>Gains</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention Focus</td>
<td>Fit</td>
<td>Mismatch</td>
</tr>
<tr>
<td></td>
<td>Mismatch</td>
<td>Fit</td>
</tr>
</tbody>
</table>

- Almost all cognitive research involves a promotion focus and a gains reward structure
  - **Promotion focus**: small monetary reward or social contract with experimenter.
  - **Gains**: reward for correct response, no reward for error
### Three-way interaction

#### Exploration optimal

<table>
<thead>
<tr>
<th>Promotion</th>
<th>Gains</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit: Good</td>
<td>Mismatch: Poor</td>
<td></td>
</tr>
<tr>
<td>Mismatch: Poor</td>
<td>Fit: Good</td>
<td></td>
</tr>
</tbody>
</table>

#### Exploitation optimal

<table>
<thead>
<tr>
<th>Promotion</th>
<th>Gains</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit: Poor</td>
<td>Mismatch: Good</td>
<td></td>
</tr>
<tr>
<td>Mismatch: Good</td>
<td>Fit: Poor</td>
<td></td>
</tr>
</tbody>
</table>
Choice/Gambling task

• Does Regulatory Fit affect choice?
• Two-Deck variant of Iowa Gambling task
  – Task 1: Exploration Optimal
  – Task 2: Exploitation Optimal
• Regulatory Focus (Global incentive)
  – Earn ticket or avoid losing ticket
• Reward Structure (Local incentive)
  – Gains vs. Losses
Gains Condition Example
PICK A CARD!

Yes
Bonus
No

450
174
0
PICK A CARD!
PICK A CARD!

Bonus
No
Yes

450
184
0
Losses Condition Example
PICK A CARD!

Yes

Bonus

No

- 450

- 174

0
Regulatory Fit and Choice

• At any moment, you have an estimate of the relative goodness of the decks
  – If you choose deterministically from the better deck, you are *exploiting*
  – If you choose more probabilistically, you are *exploring*
  – Does regulatory fit lead to more exploration than regulatory mismatch?
Modeling Choice Behavior

- EVs of each option are updated via a recency-weighted algorithm

\[ EV_{k+1} = EV_k + \alpha [ r_{k+1} - EV_k ] \]

- If reward is greater than the current EV the EV increases
- If reward is less than the current EV the EV decreases
\[ EV_{k+1} = EV_k + \alpha [ r_{k+1} - EV_k ] \]

- \( \alpha \) is a free parameter constrained to be between 0 and 1
- Higher \( \alpha \) values give greater weight to recent rewards
- When \( \alpha = 1 \), Updating Equation reduces to:
  \[ EV_{k+1} = r_{k+1} \]
- Alternatively, when \( \alpha = 0 \), Updating Equation reduces to:
  \[ EV_{k+1} = EV_k \]
Action Selection

- Action selection is probabilistically determined via choice rules (e.g. Luce, 1959)

**Softmax Rule**

\[
\begin{align*}
P_{a,t} & = \frac{e^{(\gamma EV_t(a))}}{\sum_{b=1}^{n} e^{(\gamma EV_t(b))}}
\end{align*}
\]

- Higher \( \gamma \) values indicate greater exploitation
- Lower \( \gamma \) values indicate greater exploration
Exploration Optimal

Points Based on Number Drawn From Advantageous Deck

- Predictions
  - Regulatory Fit should perform better
  - Regulatory Fit should yield small exploitation parameter (defined shortly)
Exploration optimal - Points Analysis

Average Distance from Criterion

Points Below Criterion

<table>
<thead>
<tr>
<th>Points Below Criterion</th>
<th>GAIN</th>
<th>LOSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>-40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- Promotion
- Prevention
Exploration/Exploitation parameter values

- 0.0
- 0.1
- 0.2
- 0.3
- 0.4
- 0.5
- 0.6
- 0.7

Gains Losses

Promotion
Prevention

Exploitation Parameter
(larger value = greater exploitation)
Exploitation Optimal Results (Gains only)

- Predictions supported
Summary

• Regulatory Fit $\rightarrow$ Exploratory Behavior

• Fit $\rightarrow$ Good performance when Exploration Optimal

• Fit $\rightarrow$ Poor performance when Exploitation Optimal

• Replicates pattern seen in classification (Maddox et al, 2006; Grimm et al, 2008)
Affect and Choice

- Four-Deck variant of Iowa Gambling task
  - Exploitation Optimal
- Alternative method for inducing regulatory focus
  - Smile vs. Frown faces on all cards
- Reward Structure
  - Gains vs. Losses
PICK A CARD!

<table>
<thead>
<tr>
<th>Yes Bonus</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>450</td>
<td>174</td>
</tr>
</tbody>
</table>
PICK A CARD!

Yes
Bonus
No

450

174
Predictions

• Since exploitation optimal, and assuming
  – smile = promotion
  – Frown = prevention

• Predictions
  – Regulatory Fit should perform worse
  – Regulatory Fit should yield small exploitation parameter
Points Analysis

Distance to Criterion in Points

Distance to Criterion

Gains

Losses

Positive

Negative
Exploitation Parameter

Exploitation Parameters Estimated by the Softmax Model

- Gains
  - Positive
  - Negative

- Losses
  - Positive
  - Negative
Summary

• Predictions supported
• Same behavioral and model pattern for regulatory focus and affect manipulation

• Follow-up studies (running)
  – Exploration optimal task in progress for affect task.
  – Model comparison project
  – Feedback/ITI delays
Social Motivation and Cognition

• Choice studies so far
  – Explicit incentives to induce regulatory focus
  – Affect to induce regulatory focus

• Other social factors can affect regulatory focus
  – Stereotype threat:
    • Negative self-relevant stereotype -> poor performance
    • Negative stereotypes may induce a prevention focus
    • If so, losses environment should attenuate effect.

  – DoD relevant due to hierarchical structure
Exploration Optimal Classification

\[ o = \text{category A} = \text{long, steep lines} \]
\[ + = \text{category B} = \text{all others} \]
Possible Rule-based Strategies

100% accuracy
Experiment Screen Sample

Gains

Losses
Method

• Three-dimensional classification task
  – Exploration is optimal
• Arbitrary stereotypes given to participants
  – Women are better
  – Men are better
• Manipulated gains and losses of points
• Predictions
  – Traditional stereotype threat result for gains
  – Reversed stereotype threat result for losses
Task Accuracy

Experiment 1: Women are Better

Experiment 2: Men are Better
Model-Analyses - CJ Use

Experiment 1: Women are Better

Experiment 2: Men are Better
Work In Progress

• **Exploitation optimal** task in progress
  – Involves information-integration classification
  – Prediction: Pattern should completely reverse

• End of semester effect
  – Prevention focused so better with losses
  – supported
State and Trait Factors Affect Global Incentive Focus

- Manipulate global incentive focus (state variables)
  - Explicit monetary
  - Affect/Social stereotype

- Trait variables
  - Procrastinators (end-of-semester)
  - Personality characteristics
    - Impulsivity, sensation seeking, anxiety, depression
    - IMPASS -> bias toward simple rules (Tharp, Pickering & Maddox, under review)
Task and Model Extensions
Signal Detection

- Two-stimulus identification (line length)
- Promotion/Prevention x Gains/Losses
- Biased payoffs so accuracy-maximization must be abandoned (exploration optimal)
Preliminary Results

- Early learning effect on sensitivity.
  - Fit leads to increased sensitivity.
- No systematic effects on bias.
Effect emerges on bias with extended training
  – Fit leads to bias shift toward optimal.
No systematic effects on sensitivity
Confidence paradigm

- Classification and Confidence judgment obtained
Nested Modeling Approach
(derived from Mueller & Weidmann and Maddox & Bohil)

\[ \sigma_a = \sigma_{og} = \sigma_{ol} = \sigma_{cg} = \sigma_{cl}; \beta_a = \beta_{og} = \beta_{ol} = \beta_{cg} = \beta_{cl} \]

Criterion (Focus)
[\sigma_a; \beta_o; \beta_c]

Criterion (Reward)
[\sigma_a; \beta_g; \beta_i]

Criterion (Fit)
[\sigma_o; \sigma_g; \beta_a]

Noise (Focus)
[\sigma_o; \sigma_e; \beta_o]

Noise (Reward)
[\sigma_g; \sigma_l; \beta_g]

Noise (Fit)
[\sigma_f; \sigma_m; \beta_f]

Noise (Null)
[\sigma_a; all \beta \text{ free}]

Noise (Free)
[all \sigma \text{ free}; \beta_a]
Preliminary Model Results

- **Classification Noise**: Increased classification and confidence noise, likely due to increased exploration.

- **Confidence Noise**: Increased noise in promotion prevention criteria, with gains and losses indicated.

- **Gains** and **Losses** in classification and confidence noise are shown in the graphs.
Followup

- Incorporate into Maddox and Bohil’s Hybrid Model
- External decision criterion
- ....
Summary LOCUS Plot

LOCUS Analysis Results for Motivation Experiments

- Exploration-optimal tasks
  - Strong interactions
  - No consistent main effects

- Exploitation-optimal tasks

Zhang, et al (1997; Journal of Neuroscience)
Summary

• What does it mean to “motivate” someone to do “well”?

• How do we achieve this aim?

• It is complex, but systematic and understandable.

• It involves a three-way interaction of
  – Global incentives
  – Local incentive
  – Task demand (i.e., optimal classifier strategy)
Summary (cont.)

• Regulatory Fit (interaction between global and local incentives) leads to increased exploration.

• Exploration can be advantageous or disadvantageous, depending upon the task demands.
Summary (cont.)

• We successfully applied a reinforcement learning model to choice and identified an “exploitation” parameter that tracks regulatory fit effects.

• We applied classification learning models to stereotype threat data and found that regulatory fit affects the flexibility of hypothesis-testing.

• We are extending the approach to more basic tasks such as signal detection and criterion learning and are generalizing relevant models to account for regulatory fit effects

• Finally, we are extending the approach to more dynamic decision making tasks and model development is ongoing.
Future Directions

• Continue model development
• Applications to resource acquisition (foraging)
• Exploration of other social effects on motivation
  – Social influences on choking under pressure.
Interaction with Other Groups and Organizations

• Interactions with AFOSR recipient (Brad Love)
• Interactions with the Institute for Advanced Technology (IAT) at UT-Austin, an Army UARC
• Interactions with the Institute for Innovation Creativity and Capital (IC\textsuperscript{2}) at UT-Austin
• Interactions with the Imaging Research Center (IRC) at UT-Austin
• Interactions with the Institute for Neuroscience (INS) at UT-Austin
• Interactions with the Center for Perceptual Systems (CPS) at UT-Austin
• Interactions with Veterans Affairs Medical Center (VAMC) at UC-San Diego
List of Publications Attributed to the Grant (2008-9)

Peer-Reviewed Manuscripts

• Maddox, W.T., Glass, B.D., & Markman, A.B. (under revision) Regulatory fit effects on stimulus identification.
• Grimm, L.R., Markman, A.B., & Maddox, W.T. (under revision) Regulatory fit created by time of semester and task reward structure influences test performance.
• Glass, B.C., Markman, A.B., & Maddox, W.T. (under review) The generalized exploration model (GEM): A model of human foraging for empirical analysis
• Markman, A.B., Beer, J.S., Grimm, L.R., Rein, J.R., & Maddox, W.T. (under review) The optimal level of fuzz: Case studies in a methodology for psychological research.

Conference Presentations

End-of-Semester

- End of semester participants are “bad”, “unmotivated”
- Maybe in a prevention focus?
- So mismatch with most task reward structures (gains).
- GRE math problems

Grimm, Markman, & Maddox (under review)
Regulatory Fit = Exploration: Why?

- Empirical support in several domains
- Connection to Neuroscience
  - Positive affect-frontal exploration hypothesis (Isen, Ashby, etc)
  - Regulatory focus-frontal activation findings (Amodio, Cunningham, etc)
  - LC-NE-exploration/exploitation relation (Ashton-Jones, Cohen, Daw)
Increased ITI shown to increase exploitation in an exploration optimal task (Bogacz et al, 2007)
Design and Results

- Four-deck **exploitation optimal** task (gains only)

- Increased feedback duration -> less switching, less exploration.
Risky Decisions/Feedback Interval

- Each deck has a partner
- Same EV, but one low and one high variance
- Short ITI only (gains only)

### Estimated Exploitation

<table>
<thead>
<tr>
<th>Exploitation</th>
<th>Short</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Proportion of Low Variance Responses

<table>
<thead>
<tr>
<th>Low Variance</th>
<th>Short</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Replicate effect: Increased feedback duration -> less exploration.
- Increased feedback duration -> fewer risky choices.
Followups in progress

- Losses variants
- Exploration optimal variants
Extending Models

• Choice models use one of two decision rules
  – Matching rules

\[ P_{a,t} = \frac{EV_t(a)}{EV_t(a) + EV_t(b)} \]

These rules predict that choices are affected by scalar additions to reward values, but not by scalar multiplications.

– Difference rules

\[ P_{a,t} = \frac{e^{(EV_t(a))}}{e^{(EV_t(a))} + e^{(EV_t(b))}} \]

These rules predict that choices are affected by scalar multiplications of reward values, but not by scalar additions.

Worthy, Maddox, & Markman (2008; M&C)
Testing influence of reward value

- Exp 1 (Exploitation optimal)
- Exp 2 (exploration optimal)
  - Control: Deck values 1-10
    - Deck A: EV=6; Deck B: EV=4
  - Distance-Preserving: Deck values 81-90
    - Deck A: EV=86; Deck B: 84
  - Ratio-Preserving: Deck values 10-100
    - Deck A: EV=60; Deck B: EV=40
Results

- Altering ratios has largest effect on performance
Summary

• Support for 3-way interaction in choice
  – Fit -> exploration
• Affect manipulation similar to focus
• Feedback delay increases exploitation, reduces risky choices
  – Implications for training....
• Ratio preserving models supported
Rising Optimum Task

Optimal response allocation requires escaping a local minimum of reward [taken from Bogacz et al. (2007) and Montague and Berns (2002)]

Exploration optimal

- Are people in regulatory fit less sensitive to local changes in payoffs? If they are, they will be able to overcome the local minimum

Prediction: Regulatory fit should perform better
Preliminary Results

Mean Distance From Criterion

Distance in Points

promotion-gains  prevention-gains  promotion-losses  prevention-losses

Model development in progress