# UC Merced

## Proceedings of the Annual Meeting of the Cognitive Science Society

**Title**

Computational, Neuroscientific, and Lifespan Perspectives on the Exploration-Exploitation Dilemma

**Permalink**

https://escholarship.org/uc/item/6k90f413

**Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

**ISSN**

1069-7977

**Authors**

Otto, A. Ross
Knox, Bradley
Love, Bradley
et al.

**Publication Date**

2011

Peer reviewed

# Computational, Neuroscientific, and Lifespan Perspectives on the Exploration-Exploitation Dilemma

**A. Ross Otto[1] (*Moderator*) , Bradley W. Knox[2] , and Bradley C. Love[1]**
[1] Department of Psychology, [2] Department of Computer Science, University of Texas at Austin

**Sam Gershman and Yael Niv**
Department of Psychology, Princeton University

**Darrell A. Worthy[1] and W. Todd Maddox[2]**
[1] Department of Psychology, Texas A&M University, [2]Department of Psychology, University of Texas at Austin

**Jared M. Hotaling, Jerome R. Busemeyer, and Richard M. Shiffrin**
Department of Psychology, Indiana University

**Keywords:** decision-making; computational modeling; cognitive neuroscience, information search

Consider the following real-life decisions that we make: deciding which route to take home to minimize time spent traveling, choosing amongst a set of known restaurants or a new restaurant when dining out, deciding between reading a new book by a consistently good author versus an author whose books vary widely in quality. All of these decisions involve balancing the conflicting demands of exploiting previous knowledge in order maximize payoffs versus exploring less-known options in order to gain information about the currently optimal course of action. Indeed, successfully balancing these competing demands is a non-trivial problem of interest to artificial intelligence and neural Reinforcement Learning (RL) research communities alike (Cohen, McClure, & Yu, 2007; Daw et al., 2006; Sutton & Barto, 1998). There are adverse consequences for failing to properly balance these demands in the above examples: solely making exploitative choices entails the possibility of ignorance about better courses of action, while exploring too frequently incurs large opportunity costs. The goal of the proposed symposium is to bring together researchers from a variety of perspectives who are working to better understand the psychological and neurobiological mechanisms underlying exploratory choice.

In recent years, novel computational modeling approaches have been developed and applied to understanding how humans incorporate the demands of information gathering into their patterns of choice. These modeling techniques have yielded insight not only in describing human choice behavior, but also in understanding the neurobiological and physiological correlates of exploratory decision-making in humans (Daw et al., 2006; Jepma & Niewenhuis, in press). The researchers who have agreed to participate in this symposium are all applying computational models to better understand the psychological and neurobiological mechanisms underpinning peoples negotiation of the exploration-exploitation tradeoff. The modeling approaches taken by these speakers are indeed diverse, ranging from uncovering hidden variables underlying decision-makers' choices in order to unpack neu-robiological and physiological measurements to describing aging-related changes in exploratory choice behavior. The proposed symposium will provide a forum for highlighting recent advances in applications of computational modeling to human exploratory choice.

The speakers who have agreed to participate in this symposium–while each performing research that elucidates psychological and neurobiological mechanisms underlying exploratory decision-making—offer different perspectives on the issue. The research described includes 1) aging work examining lifespan changes in exploratory decision-making, elucidating the underlying neurobiology of these types of choices (Worthy & Maddox), 2) a Bayesian account of effect of novelty–when humans are presented with new, potentially rewarding options–on exploratory choice, and how these novelty signals are represented in the brain in order to compute values and guide choices (Gershman & Niv), 3) how individuals incorporate uncertainty and information search costs when planning courses of action in situations with sequential dependencies between choices and outcomes (Hotaling, Busemeyer, & Shiffrin), and 4) how internally calculated uncertainty about the environment directs exploratory choice and manifests itself physiologically over the course of decision-making (Otto, Knox, & Love). In addition to proposing answers to a diverse set of important psychological and neuroscientific questions, the lines of research described by these speakers rely upon laboratory tasks with monetary incentives that, each in their own way, incorporate ecologically interesting choice and reward dynamics.

## Belief-directed Exploration in Human Decision-Makers: Behavioral and Physiological Evidence  *A Ross Otto, W. Bradley Knox, & Bradley C. Love*

Decision-making in uncertain environments poses a conflict between the goals of exploiting past knowledge in order to maximize rewards and exploring less-known options in order to gather information. However, the descriptive modeling framework utilized in previous studies of exploratory choice behavior characterizes exploration as the result of choices, rather than a process reflecting beliefs and/or un-

certainty about the environment. This work presents an Ideal Actor model that prescribes an optimal incremental belief-update procedure and payoffs-maximizing pattern of choice in a novel decision-making task. By comparing human choice dynamics to those prescribed by the Ideal Observer/Actor, I evaluate the notion that people conduct exploration in a belief-directed fashion. Further, I reveal how hidden, internally calculated beliefs are indexed by pre-choice autonomic arousal (measured using skin conductance) and choice reaction times in the course of decision-making, providing evidence that people indeed negotiate the exploration-exploitation choice tradeoff in a belief-directed fashion. This model-based analysis provides a quantitative analysis of how uncertainty relates to anticipatory autonomic arousal preceding choices will elucidate the role of autonomic arousal in risky decision-making, a topic of much debate in the literature.

**Reinforcement Learning, Exploration, and Novelty Bonuses in the Brain** *Sam Gershman & Yael Niv*

The "puzzle of novelty" refers to the contradictory observations that humans and animals are both neophobic (repulsed by novelty) and neophilic (attracted to novelty). In an effort to shed computational light on this puzzle, we have analyzed neotic preferences in terms of Reinforcement Learning (RL) models. We show mathematically that both neophilia and neophobia can arise from Bayesian inductive generalization over reward predictions, depending on the reinforcement history. In essence, the Bayesian RL model regularizes reward predictions for novel actions towards the predictions for other actions taken in the same context. This model offers a new perspective on so-called "novelty bonuses"optimistic reward predictions that encourage exploration. According to the Bayesian RL model, the degree of bonus (or penalty) will depend on experience with other actions in the same context. We present behavioral and brain imaging evidence consistent with this model, showing that neural valuation signals are better described by the Bayesian model than by a novelty bonus model. These findings suggest that human exploratory tendencies are guided by inductive knowledge about the environment.

**Effects of Normal Aging on the Resolution of the Exploration/Exploitation Dilemma in Decision-Making** *Darrell A. Worthy & W. Todd Maddox*

We examined how normal aging affects predispositions toward either exploratory or exploitative choice by having healthy younger and older adults perform two decision-making tasks where exploitation (Experiment 1) or exploration (Experiment 2) was the optimal strategy. In both experiments participants performed two variants of the task, one week apart, where they had to either maximize points gained or minimize points lost. A reinforcement learning model was used to directly parameterize the degree to which subjects exploited the options with the largest expected rewards versus explored options with lower expected rewards. Older adults

engaged in more exploration across both tasks, performing worse in Experiment 1, but better in Experiment 2. Their data were also fit best by lower exploitation parameter values relative to the data of younger adults. Two additional experiments were conducted that required the learning of a hidden, dynamic structure of the reward environment. It was hypothesized that engaging in a more systematic form of exploration would aide in uncovering the underlying structure of the environment. Older adults performed more optimally than younger adults in both of these experiments. We conclude by discussing possible behavioral and neurobiological reasons for these age-based differences along the exploration/exploitation continuum.

**Information Search in Multi-Stage Risky Decision-Making** *Jared M. Hotaling, Jerome B. Busemeyer, & Richard M. Shiffrin*

Research into risky decision-making has traditionally presented individuals with choice alternatives that provide an immediate reward or punishment based on the outcome of a single random event. Decisions are typically made in isolation, independent from any previous or subsequent choices. This approach neglects the complexity of everyday decision-making, which often involves multiple interdependent choices and several uncertain events. We present recent work that extends the traditional risky decision making paradigm by incorporating some of the complexities of real world choices. Participants completed a series of multistage decision trials, represented as branching decision trees. At decision nodes, participants chose which path to take through the tree. At chance nodes, a random event determined the path. Crucially, participants had the option to use some of the points earned on previous trials to reduce their uncertainty by purchasing information about chance nodes. We review data showing how individuals incorporate factors like risk, information search cost, and degree of uncertainty when forming plans for multistage decision scenarios. Our results show individual differences, with several distinct strategies emerging. A comparison of multiple competing models is used to elucidate the cognitive processes at work.

## References

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007, May). Should i stay or should i go? how the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933 –942.

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

Jepma, M., & Nieuwenhuis, S. (n.d.). Pupil diameter predicts changes in the ExplorationExploitation tradeoff: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, 1–10. (in press)

Sutton, R., & Barto, A. G. (1998). *Reinforcement learning*. Cambridge, MA: MIT Press.