Symposium: Dynamic Decision Making

Todd M. Gureckis (*Moderator*) and Douglas Markant
Department of Psychology, New York University

Jared M. Hotaling, Eric Dimperio, and Jerome R. Busemeyer
Department of Psychological and Brain Sciences, Indiana University

Michael D. Lee, Shunan Zhang, and Mark Steyvers
Department of Cognitive Sciences, University of California at Irvine

Bradley C. Love and A. Ross Otto
Department of Psychology, University of Texas at Austin

Dylan A. Simon and Nathaniel Daw
Center for Neural Science, New York University

Keywords: sequential decision making, computational modeling, cognitive neuroscience

The experimental study of decision making has historically focused on simple single-trial judgement or reasoning tasks. However, real world behavior often necessitates online decision making, planning, and sequentially organized behavior. The goal of the proposed symposium is to bring together researchers who are working to understand the cognitive processes underlying this kind of "dynamic decision making" (defined as tasks or contexts that are structured as a sequence of interdependent decisions).

A symposium on this topic is particularly timely since research in the area of dynamic decision making is having a tremendous impact on the field of psychology as a whole. First, researchers are converging on a set of novel computational modeling approaches that explain how decision makers plan sequences of multiple actions, take into account future contingencies, and react in real time to continually changing environmental dynamics. Second, many of the proposed algorithms and models are closely linked to neurobiological correlates (e.g., the recent explosion of research on neurobiology of reinforcement learning). Third, many of the tasks that are being developed for evaluating these models also appear to relate to important individual differences in real-world decision-making. The goal in the symposium is to 1) highlight some of the best work in this area, 2) to facilitate communication between researchers working on these problems from varying perspectives, and 3) to provide an excellent showcase of this area for members of the cognitive science community who may not yet be familiar with this work.

The speakers who agreed to participate are all accomplished researchers in this area but each approach the set of problems involved in sequential decision making and learning from a slightly different perspective. The key topics covered include 1) how people plan sequences of actions to accomplish goals (Hotaling, Dimperio, & Busemeyer, Simon & Daw), 2) the underlying neurobiology of sequential decision making and planning (Simon & Daw), 3) how cognitive representations of the task or environment supporting planning and decision-making (Gureckis & Markant, Love & Otto, and Simon & Daw), and 4) how people balance exploration and exploitation in order to arrive at effective decision strategies in an unknown environment (Lee, Zhang, and Steyvers and Gureckis & Markant). In addition to these overlapping psychological themes, the researchers all share a core approach of applying sophisticated computational models to understand human behavior (including Bayesian approaches, reinforcement learning, and Markov Decision Processes).

Todd Gureckis & Douglas Markant (New York University)

Exploring to Exploit: Modeling the Process of Information Search and Planning

Effective learning often involves actively querying the environment for information that can be exploited at a later point in time. However, the space of observations available in any situation can vary greatly in potential "informativeness" and relative cost. How do people decide which observations to make at any point in time given their future goals? We describe a series of studies looking at how people plan sequences of information collection actions in a cognitive search task based on the children's game Battleship. Participants made sequences of observations to disambiguate between a large number of potential game configurations subject to information-collection costs. Computational models are developed which predict which observations people will make on any given trial and when they should stop collecting information and exploit their current knowledge. In particular, the models measure the degree to which individuals take into account future consequences when planning immediate actions. In our second study, we explore how people generate hypotheses consistent with their prior beliefs and how these hypotheses in turn influence search behavior.

Jared M. Hotaling, Eric Dimperio, & Jerome R. Busemeyer (Indiana University)

Cognitive Models of Planning Behavior in Multi-Stage Risky Decision Making

Much 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 random event. This allows researchers to understand how the values of choice alternatives and the probabilities associated with risk can influence an individuals choices. We present recent work that extends that research by manipulating the outcome probabilities and rewards involved in a multistage decision task where some rewards are only possible after a sequence of decisions. Our results show individual differences, with some participants being sensitive to possible future rewards and likelihoods, and others appearing not to plan ahead. A comparison of multiple competing models helps identify decision processes when planning ahead did occur.

Michael Lee, Shunan Zhang, & Mark Steyvers (Univ. of California, Irvine)

Human and optimal exploration and exploitation in sequential decision-making

In bandit problems, a decision-maker chooses repeatedly between a set of alternatives. They get feedback after every decision, either recording a reward or a failure. They also know that each alternative has some fixed unknown probability of providing a reward when it is chosen. The goal of the decision-maker is to obtain the maximum number of rewards over all the trials they complete. Bandit problems provide an interesting formal setting for studying the balance between exploration and exploitation in decision-making. In early trials, it makes sense to explore different alternatives, searching for those with the highest reward rates. In later trials, it makes sense to exploit those alternatives known to be good, by choosing them repeatedly. How exactly this balance between exploration and exploitation should be managed, and should be influenced by factors such as the distribution of reward rates, the total number of trials, and so on, raises basic questions about adaptation, planning, and learning in intelligent systems. In this talk, we present a series of models, both Bayesian and heuristic, aimed at understanding how people balance exploration and exploitation, and how their strategies relate to optimal decision-making.

Brad Love & A. Ross Otto (University of Texas at Austin)

You Don't Want To Know What You're Missing: When Information about Forgone Rewards Impedes Dynamic Decision Making

When learning to make decisions from experience, one reasonable intuition is that adding relevant information should improve performance. In contrast, we find that additional information about foregone rewards (i.e., what could have gained at each point by making a different choice) severely hinders participants ability to repeatedly make choices that maximize long-term gains. We conclude that foregone reward information accentuates the local superiority of short-term options (e.g., consumption) and consequently bias choice away from productive long-term options (e.g., exercise). These conclusions are anticipated by a standard reinforcement learning mechanism that processes information about experienced and forgone rewards. In contrast to related contributions using delay-of-gratification paradigms, we do not posit separate top-down and emotion-driven systems to explain performance. We find that individual and group data are well characterized by a single reinforcement learning mechanism that combines information about experienced and foregone rewards. These findings will be situated within a broader research program that aims to characterize how people explore and exploit environments with unknown rewards and poorly understood states. Finally, interventions for improving human performance will be discussed.

Dylan Simon & Nathaniel Daw (New York University)

Neural correlates of decision evaluation by forward planning in sequential tasks

Theoretical models of reinforcement learning are commonly applied within neuroscience to explain the neural processes involved in learning and decision making. However, the approaches used are predominately "model-free," such as temporal difference learning which learns action values or policies directly from reinforcement without explicitly representing or utilizing any information about task structure. While these theories have shed light on observed neural activity in simple "bandit" tasks involving repeated choices rewarded independently and immediately, it is at odds with a long line of behavioral evidence from psychology and cognitive science for more flexible, goal-directed forward planning processes. We show how a different set of "model-based" reinforcement learning algorithms can be used to account for these phenomena, and test this framework in humans using a number of dynamic, continuous, sequential decision tasks. The models can account both for observed choice behavior and fMRI BOLD signals in decision-related brain areas. Consistent with cognitive theories, both behavior and neural activity show evidence of flexible learning and forward planning, indicating that existing neural models provide an incomplete picture of learning and decision making in dynamic tasks.