Exploration strategies in human decision making

Robert C. Wilson, Andra Geana, John M. White, Elliot A. Ludvig and Jonathan D. Cohen

Abstract (100 words)

The tradeoff between pursuing a known reward (exploitation) and sampling unknown, potentially better opportunities (exploration) is a fundamental decision-making challenge. Theories formalize the value of exploration as an information bonus, driving directed exploration. However, this is difficult to compute; a simpler alternative is to increase decision noise, driving random exploration. We designed a task to quantify these strategies. As the opportunity to explore increased so did the information bonus and the decision noise. This suggests that humans use and adapt both an information bonus and decision noise to solve explore-exploit problems in practice.

Body (600 words)

When you go to your favorite restaurant do you always order the same thing, or do you sometimes try something new? Sticking with an old favorite ensures a good meal, but if you are willing to explore you might discover something better – or something worse. This simple conundrum, deciding between something you know about or trying something new and unknown, is referred to as the exploration-exploitation dilemma [1, 2]. Whether deciding on a meal, a career, or a life partner, this is an important and recurrent problem at all levels of decision-making.

Theoretical accounts suggest two distinct strategies for resolving this dilemma. One is directed exploration [3], in which choices are biased towards ambiguous (and hence more informative) options with an `information bonus'. The other strategy is random exploration [4], in which choices are biased by internal decision noise. Directed strategies derive from theories of optimal decision-making that ensure the greatest amount of reward in the long run, whereas random strategies reflect simpler heuristics that may be less costly to implement in practice.

There has been growing interest in a number of fields (economics, psychology and neuroscience) in how humans solve the explore-exploit dilemma, but to date, relatively few have dissected the strategies involved. Results on directed exploration are mixed – with some studies finding it (e.g. [5]) and others not (e.g. [6]) – while (to the best of our knowledge) no one has studied random exploration at all. We believe that this lack of progress reflects the subtle complications that arise in explore-exploit experiments. In particular, directed exploration is difficult to measure, because ambiguity aversion acts counter to the information bonus and because sequential choice tasks subtly confound reward and information – as subjects choose more rewarding options more often, over time they are better informed about options with higher rewards. Random exploration is also hard to

assess because noise is hard to measure: how do we distinguish randomness in behavior from misspecification of the behavioral model?

To address these limitations, we examined decision-making behavior in a simple task in which participants were given prior information about each of two options and then allowed to make a series of choices between them. We experimentally manipulated the amount of information that participants were given about each option (i.e., the ambiguity of each), as well as the number of choices they would be allowed to make between those options (i.e. the game horizon). By controlling the information subjects received we removed the confound between reward and information. By varying the horizon we could dissociate baseline levels of decision noise and ambiguity aversion from exploration induced changes.

We quantified our analysis by fitting a formal model of the decision-making process to participants' behavior. This approach allowed us to determine not only the overall amount of exploratory behavior, but also dissociate the influence of the two strategies: information seeking (directed exploration) and decision noise (random exploration). We found evidence for both types of exploratory behavior and furthermore that these strategies were adaptively modulated by the opportunity to explore. The result for decision noise is especially surprising as it suggests that the levels of random decision noise can be utilized and controlled as a means of exploration.

[1] Kaelbling et al. Journal of Artificial Intelligence Research, 4, 1996.
[2] Sutton and Barto. Reinforcement learning: An introduction. MIT Press, 1998.
[3] Gittins and Jones. In J. Gans, editor, Progress in statistics, pages 241–266. Amsterdam, The Netherlands: North-Holland, 1974.
[4] Luce. Individual Choice Behavior. Wiley, NY., 1959.
[5] Meyer and Shi. Management Science, 41(5):817–834, 1995.
[6] Daw et al. Nature, 441(7095):876–9, Jun 2006.