Does a present bias influence exploratory choice?

Alexander S. Rich Todd M. Gureckis

Does a present bias influence exploratory choice?

Exploration outside the lab

Exploration has been studied outside the lab in a wide range of contexts. While these domains vary greatly in their superficial characteristics, a bias towards under-exploration has often been observed.

Learned helplessness, a phenomenon applicable to many behaviors and domains, has been described as an example of insufficient exploration. In learned helplessness, an organism experiences the absence of control over the environment, learns that the environment is uncontrollable, and thus ceases to take actions that might allow it to discover that it can in fact exert control. Learned helplessness has been proposed to underly some forms of depression (Lyn Y Abramson, Metalsky, & Alloy, 1989; L Y Abramson, Seligman, & Teasdale, 1978) as well as problems ranging from difficulties in school (Diener & Dweck, 1978) to poverty (Evans, Gonnella, Marcynyszyn, Gentile, & Salpekar, 2005). While the cognitive appraisal of experienced events affects the development of learned helplessness (L Y Abramson et al., 1978), patterns of exploration clearly play a role as well (Huys & Dayan, 2009, Teodorescu & Erev (2014a)). In the case of depression, interventions aimed at increasing the exploration of activities that might be rewarding have been found to be as effective as those with a more cognitive orientation (Jacobson et al., 1996).

Under-exploration also seems to occur in the development of complex skills, such as flying a plane or playing a sport (D. Gopher, Weil, & Siegel, 1989). In these settings, an "emphasis change" training method that encourages people to continually explore the performance space leads to greater performance gains than unguided practice or more complex training methods. Without this intervention, people often enter a "local maximum" in which exploration decreases and performance plateaus (Yechiam, Erev, & Gopher, 2001).

In many other areas under-exploration is less clearly established, but is suspected to play a role in maladaptive behavior. Insufficient exploratory interaction with outgroups may be one cause of stereotypes and prejudice (Denrell, 2005), and interventions that increase inter-group contact reduce stereotypes (Shook & Fazio, 2008). The crowding out of exploration by exploitation is a concern in organizational behavior as well (Levinthal & March, 1993; March, 1991), prompting research into organizational structures that may preserve exploration (Fang, Lee, & Schilling, 2010).

Exploration in the lab

Lab studies of exploratory choice have allowed researchers to fully control the reward structure of the environment and precisely measure behavior, as well as comparing behavior to optimal patterns of choice and other formal models. These studies have yielded a number of insights into the factors leading to more or less exploration, including aspiration levels (**???**), uncertainty (Speekenbrink & Konstantinidis, 2015), and the future value of information (Rich & Gureckis, 2017; Wilson, Geana, White, Ludvig, & Cohen, 2014)

Interestingly, under-exploration has not emerged as a clear pattern in lab experiments. Instead, results are mixed with people sometimes under-exploring, sometimes over-exploring, and sometimes exploring close to an optimal amount. To take two illustrative examples, (Zwick, Rapoport, Lo, & Muthukrishnan, 2003) found that in a sequential search task people under-searched when there were no information costs but over-searched when there were information costs, and (Teodorescu & Erev, 2014b) found that people explored unknown alternatives too often or not often enough depending on whether rare outcomes were positive or negative. Similar results have been obtained within and across a variety of other studies and paradigms (Hertwig, Barron, Weber, & Erev, 2004; Juni, Gureckis, & Maloney, 2016; Navarro, Newell, & Schulze, 2016; Sang, Todd, & Goldstone, 2011; Tversky & Edwards, 1966).

Exploration and myopia

Formal analysis

Capturing present bias in the lab

Experiment

Methods

Participants.

Design and procedure.

Barratt Impulsiveness Scale

Prior to reading the experiment instructions, participants completed the 30-item Barratt Impulsiveness scale (Patton, Stanford, & Barratt, 1995) on the computer.

Consumption tasks.

Participants were informed that there were two types of tasks, a “slider task” and a “video task,” that they would complete during 30-second “work periods.” The number of remaining work periods in the experiment was shown at the top of the screen, as was the number of seconds left in the current work period.

The slider task was based of a task previously used by (Gill & Prowse, 2012). In each period of the slider task, five horizontal sliders appeared on the screen (Figure XXX). Each started at a random setting between 0 and 100, with the slider’s value shown to its right, and with a random horizontal offset so that the sliders were not aligned. The participant’s task was to use the mouse to move each slider to “50” before the work period ended. When a participant released the mouse at the correct setting, the slider turned green to show it had been completed. To ensure that the task took close to the allotted 30 seconds, at the beginning of the task only the top slider was enabled, and the other four were grayed out. Additional sliders were enabled at five-second intervals, such that all five sliders were available after 20 seconds. To make the slider task more unpleasant, a short static noise was played through the computer speakers intermittently at a moderate volume during the task.

The video tasks consisted of simply watching one of four videos: an episode of “Planet Earth”, and episode of “The Great British Bakeoff”, and episode of “Unchained Reactions”, or an Ellen Degeneres comedy special. Participants watched the video through a player on the computer screen. They were free to fast forward or rewind the video at will, and could also switch among the videos at any time by clicking one of four tabs above the player.

To incentivize participants to attend to and perform the slider task, they were penalized if they missed more than 10% of the sliders. For each percentage over 10% of sliders that were not set to 50 over the course of the experiment, $.20 was deducted from a bonus that started at $5.00. For example, if a participant failed to set the sliders to 50 for 18% of sliders, their final bonus would be $3.40.

Choice task.

Participants completed a total of 56 work periods. The first eight were automatically spent performing the slider task. For the remaining 48, participants made a choice prior to each work period that determined whether the work period would be devoted to the slider task or the video task.

For the choice task, participants were shown a "machine" that could create slider or video tasks. The machine consisted of a black-and-gold "best" spinner and a panel of possible new spinners. Participants selected either "run best spinner" or "run new spinner".

If the participant selected "run best spinner", the spinner would visually rotate on the screen. If it landed on gold, the machine created a video task; if it landed on black, the machine created a slider task. Thus, the probability of producing a video task was equal to the proportion gold of the spinner.

If the participant selected "run new spinner", the new spinners in the panel were covered up and randomly shuffled. The participant then clicked one of the gray squares, revealing the new spinner underneath. As was explained to the particpants, and is visually apparent, one third of the possible new spinners are completely black, while the remaining two thirds range from 5% to 100% gold, in even increments of 5%.

After revealing a new spinner, it was spun, producing a video or slider task in the same manner as the "best spinner". Critically, if the new spinner selected had a higher proportion gold than the best spinner, it would replace the best spinner for future choices. Thus, choosing "run new spinner" was an exploratory action that could lead to the discovery of a better option that could be exploited in later choices.

Participants were also informed that after every work period there was a one in six chance that the machine would reset itself. In fact, the experiment was designed so that there was exactly one reset in every set of 6 trials (i.e., trials 1–6, 7–12, etc.). When the machine reset, the "best spinner" was set to a new starting value. The starting values following resets (including the intial starting value) were {20%, 25%, … 55%, 60%}, randomly ordered.

Immediate and delayed conditions.

Participants were pseudorandomly assigned to one of two conditions. In the Immediate condition, participants completed the task produced by a choice in the work period immediately following the choice. In the Delayed condition, participants completed the task produced by a choice after eight intervening work periods had passed, which was about five minutes after making the choice. This means that participants in the Delayed condition began making choices during the initial eight slider task work periods, in order to have outcomes determined when they reached the ninth and later work periods.

To make this delay intuitive, participants were shown a work queue at the bottom of the screen that contained eight tasks. In the Delayed condition, upon making a choice a new slider or video task icon was added to the right of the queue, and then the leftmost task on the queue was performed and removed. In the Immediate condition, participants were still shown the cue, but upon adding an icon to the right of the queue that outcome was performed immediately. This means that in the Immediate condition the queue acted simply as a history of the past eight outcomes.

Post-task questions.

Following the final work period, participants were asked to rate their enjoyment of the slider task and of the video task on a scale from 1 to 7, where 1 indicated extremely unenjoyable and 7 indicated extremely enjoyable.

Results

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness Depression : A Theory-Based Subtype of Depression. *Psychological Review*, *96*(2), 358–372.

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: critique and reformulation. *Journal of Abnormal Psychology*, *87*(1), 49–74. http://doi.org/10.1037/0021-843X.87.1.49

Denrell, J. (2005). Why most people disapprove of me: experience sampling in impression formation. *Psychological Review*, *112*(4), 951–978.

Diener, C. I., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, *36*(5), 451–462. http://doi.org/10.1037/0022-3514.36.5.451

Evans, G. W., Gonnella, C., Marcynyszyn, L. a, Gentile, L., & Salpekar, N. (2005). The Role of Chaos in Poverty and Children's Socioemotional Adjustment.

Psychological Science, *16*(7), 560–565.

Fang, C., Lee, J., & Schilling, M. a. (2010). Balancing Exploration and Exploitation Through Structural Design: The Isolation of Subgroups and Organizational Learning. *Organization Science*, *21*(3), 625–642. http://doi.org/10.1287/orsc.1090.0468

Gill, D., & Prowse, V. (2012). A Structural Analysis of Disappointment Aversion in a Real Effort Competition. *The American Economic Review*, *102*(1), 469–503.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, *71*(1-3), 147–177. http://doi.org/10.1016/0001-6918(89)90007-3

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–539.

Huys, Q. J. M., & Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition*, *113*(3), 314–328. http://doi.org/10.1016/j.cognition.2009.01.008

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., … Prince, S. E. (1996). A Component Analysis of Cognitive-Behavioral Treatment for Depression. *Journal of Consulting and Clinical Psychology*, *64*(2), 295–304.

Juni, M. Z., Gureckis, T. M., & Maloney, L. T. (2016). Information Sampling Behavior With Explicit Sampling Costs. *Decision*.

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, *14*, 95–112. http://doi.org/10.1002/smj.4250141009

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87.

Navarro, D. J., Newell, B. R., & Schulze, C. (2016). Learning and choosing in an uncertain world : An investigation of the explore-exploit dilemma in static and dynamic environments. *Cognitive Psychology*, *85*, 43–77. http://doi.org/10.1016/j.cogpsych.2016.01.001

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774. http://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Rich, A. S., & Gureckis, T. M. (2017). Exploratory choice reflects the future value of information. *Decision*, *(in press)*.

Sang, K., Todd, P. M., & Goldstone, R. L. (2011). Learning near-optimal search in a minimal explore/exploit task. *Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society*, 2800–2805.

Shook, N. J., & Fazio, R. H. (2008). Interracial roommate relationships: An experimental field test of the contact hypothesis: Research article. *Psychological Science*, *19*(7), 717–723. http://doi.org/10.1111/j.1467-9280.2008.02147.x

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit task. *Topics in Cognitive Science*, *7*. http://doi.org/10.1111/tops.12145

Teodorescu, K., & Erev, I. (2014a). Learned Helplessness and Learned Prevalence: Exploring the Causal Relations Among Perceived Controllability, Reward Prevalence, and Exploration. *Psychological Science*, *25*(10), 1861–1869. http://doi.org/10.1177/0956797614543022

Teodorescu, K., & Erev, I. (2014b). On the Decision to Explore New Alternatives: The Coexistence of Under- and Over-exploration. *Journal of Behavioral Decision Making*, *27*, ‘. http://doi.org/10.1002/bdm

Tversky, A., & Edwards, W. (1966). Information versus reward in binary choices. *Journal of Experimental Psychology*, *71*(5), 680.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans Use Directed and Random Exploration to Solve the Explore – Exploit Dilemma. *Journal of Experimental Psychology: General*.

Yechiam, E., Erev, I., & Gopher, D. (2001). On the potential value and limitations of emphasis change and other exploration-enhancing training methods. *Journal of Experimental Psychology. Applied*, *7*(4), 277–285. http://doi.org/10.1037/1076-898X.7.4.277

Zwick, R., Rapoport, A., Lo, A. K. C., & Muthukrishnan, a. V. (2003). Consumer Sequential Search: Not Enough or Too Much? *Marketing Science*, *22*(4), 503–519. http://doi.org/10.1287/mksc.22.4.503.24909