Foraging for Alternatives: Ecological Rationality in Keeping Options Viable

Hansjörg Neth, Neele Engelmann & Ralf Mayrhofer
{hneth|neele.engelmann|rmayrho}@uni-goettingen.de
Cognitive and Decision Sciences
Georg-Elias-Müller Institute of Psychology
University of Göttingen, Germany

Abstract

Do we invest irrational amounts of effort into keeping options viable, or do we manage available and threatened options in an adaptive fashion? To ask and answer this question, we advocate an approach that considers the dynamic properties of decision environments. By linking the exploration-vs.-exploitation dilemma to animal foraging, we show that preserving and abandoning options can both be adaptive. Specifically, people should stay and abandon options in progressive environments, and leave and seek alternatives in exhaustive environments. We extend a multi-arm bandit problem with threatened options by a manipulation of environmental expectations. Our findings show that people are highly sensitive to environmental assumptions and small payoff differentials. This replicates the original effect, but may explain the apparently irrational tendency to keep options open as an ecologically rational adaptation.

Keywords: exploration vs. exploitation; multi-arm bandit; multi-tasking; switch costs; animal foraging.

Should I stay or should I go now?
If I go there will be trouble,
and if I stay it will be double…
The Clash

Introduction

Having a choice between multiple options confronts us with a dilemma: We desire the availability of options, but keeping options open often comes with costs. For instance, the simultaneous pursuit of several careers, research goals, or romantic relationships can reduce the success or satisfaction enjoyed by any particular one. The dilemma between freedom of choice and the costs of preserving options is aggravated by the fact that options tend to deteriorate or disappear when we fail to invest enough effort or resources (e.g., attention, time, or money) into their availability.

In an intriguing study, Shin and Ariely (2004) asked whether options are valued differently when they are in danger of disappearing. The availability of multiple options was operationalized by playing a virtual "door game" in which participants first chose one of three doors to enter a room. From then on, participants could either earn a monetary reward by pressing a button within the chosen room or select a different door to switch to another room. As the overall budget of actions was limited to 100 clicks, switching between rooms incurred the opportunity cost of forgoing one payoff in the current room. In one condition, the three alternative rooms were continuously available. A second condition imposed an additional threat of disappearing options by gradually shrinking unused doors. On any click within a room or on the door to another room the two unchosen doors would lose 1/15th of their original width. This implies that the access to a room would effectively vanish if it has not been chosen within the 15 most recent clicks.

Shin and Ariely's (2004) main result was that the threat of losing options caused participants to invest additional efforts into their availability. Crucially, the costs of keeping doors open exceeded the potential benefits conveyed by them. Thus, people exhibited a strong and seemingly irrational tendency to keep poor options viable — a phenomenon that the authors attributed to an aversion to the loss of options.

An analysis of the original study reveals an extremely demanding task even when options are constantly available. Sequential decision environments involving payoff functions that must be estimated on the basis of limited experience transcend the realms of risky choice and require decision making under uncertainty (Sims, Neth, Jacobs, & Gray, 2013). Formally, the simultaneous availability of three rooms constitutes a multi-armed bandit problem (Banks, Olson, & Porter, 1997) in which participants must trade-off the exploration of unfamiliar choices with the exploitation of familiar ones to maximize the gain from probabilistic payoff functions (see Lee, Zhang, Munro, & Steyvers, 2011; Steyvers, Lee, & Wagenmakers, 2009; Sutton & Barto, 1998, for formal models). In the present paradigm, this dilemma is aggravated by implicit but substantial switch costs that are paid by forgoing one payoff whenever entering another room. A further difficulty of the original study was that the payoffs of all three rooms were sampled randomly from extremely similar payoff distributions. In fact, for the first 3 of 4 experiments the options' payoffs differed merely in their distributional form and type, but were equal in expected value. By contrast, the availability of three separate doors suggested that people actually had a choice. If participants erroneously expected that the options differed in some relevant aspect, the high similarity and overlap of experienced payoffs rendered the exploration task even more demanding. Thus, the threat of losing options was added to a non-trivial array of challenging task components.

In a follow-up study that manipulated the difference between the payoff distributions of nine options, Ejova, Navarro, and Perfors (2009) showed that people are more willing to walk away from inferior options when the discriminability between options increases. This suggests that the dysfunctional aversion to option loss occurs mostly in situations that are difficult to explore, whereas people are willing and able to sacrifice options when superior ones are available and identifiable.

An Ecological Approach

The main intuition behind our approach is that it sometimes makes perfect sense to keep options open. Precisely when this is the case depends not only on the number and discriminability of options, but also on the environment in which an organism is pursuing its goals. A fundamental axiom of cognitive science states that the rationality of an organism's behavior can only be evaluated in relation to the goals and environment in which it occurs (Simon, 1956). Adaptive behavior is characterized by the degree of fit between an organism's goals, strategies, and the structure of the environment. Behavior exhibiting this fit is called *ecologically rational* (Todd, Gigerenzer, & the ABC Research Group, 2012).

The dilemma between the benefits and costs of having choices is not restricted to human life. In fact, nature routinely confronts organisms with options that are in danger of disappearing if they are not seized in time. For instance, foraging animals may have to consume or defend some patch of food against competitors, or harvest some resource before it deteriorates or withers away.

A crucial link between the need to keep options open and the structure of environments is provided by *foraging theory* (FT, Stephens & Krebs, 1986). FT studies animal behavior as an economic resource allocation problem and assumes that animals inhabit a patchy environment in which resource gains within patches must be weighed against the costs of energy expenditure by moving between patches. The analogy between animal foraging and human cognition has been productive in modeling external information search (e.g., in libraries or on the web, Pirolli & Card, 1999), discretionary interleaving between tasks (Payne, Duggan, & Neth, 2007), and internal search processes like free recall of items from memory (Hills, Jones, & Todd, 2012).

To study the threat of vanishing options as a function of environmental characteristics, we distinguish between two basic types of environments:[1]

1. In *exhaustive* environments, patches deteriorate with use as some resource is consumed and depleted. When not attended to, resources remain stable or regenerate.
2. In *progressive* environments, patches improve with use as some resource is developed and enriched. When not attended to, resources remain stable or deteriorate.

As the energetic quality of patches declines as food is being harvested, FT mostly considers exhaustive environments. However, many options in human life tend to grow better or more satisfying with additional investments, i.e., they are of the progressive type. For instance, consider tasks that improve with practice (like playing a musical instrument), the acquisition of new skills (learning a language), or deepening social ties (fostering friend- and relationships).

[1] A complete taxonomy would need to consider the functional forms of the dynamics of a patch's usage and non-usage, as well as the causes and triggers of resource depletion and recovery.

In this paper, we extend and qualify previous results on the willingness to preserve and abandon options by investigating the impact of different environments on such decisions. Specifically, investing large efforts into the preservation of multiple options only seems dysfunctional in progressive environments, as it prevents any individual option from reaching a high level. By contrast, abandoning options only seems dysfunctional in exhaustive environments, as any depletion of a current option increases the need for alternatives. Expressed positively, our distinction provides a tentative answer to the introductory question by *The Clash*: Staying with the current option (and abandoning alternative options) is adaptive in progressive environments, and going (and pursuing alternative options) is adaptive in exhaustive environments. Consequently, ecologically rational organisms should invest more effort into preserving options in exhaustive than in progressive environments.

Experiment

Do people invest irrational amounts of effort into keeping options viable, or do they manage available and threatened options adaptively? To examine people's investments into threatened options as a function of environmental factors, we manipulated their expectations about the environment.

Participants 63 students (44 female; mean age of 23.7 years) of the University of Göttingen, Germany, participated in this experiment and either received course credit or were paid € 8 (approx. US-$ 10.80) per hour. As two participants did not engage in any exploration behavior in the last two environments, they were excluded from all analyses.

Apparatus Participants were seated individually in front of a 19"-screen and played a virtual door game that was implemented in MS PowerPoint™ with Visual Basic extensions.

To remedy some of the difficulties mentioned above and render an abstract task more intuitive and engaging, we couched our study in a Las Vegas narrative involving different types of casinos. Each casino contained three rooms in which a slot machine could be played to gain points (see Figure 1). Within a casino, participants' goal was to gain as many points as possible with 100 clicks. This budget constraint was operationalized by an initial allocation of 100 jetons. Both playing on the current room's slot machine and switching to a different room incurred a cost of one jeton. To make the costs of playing and switching equally salient, every click caused one jeton to roll off the stack of remaining jetons.

Design Our study used a mixed 2×3 design that included Shin and Ariely's (2004) original threat to the availability of options as a *between-subjects* manipulation. In a *constant* condition, all three rooms were always available. In the *threatened* condition, participants were instructed that a casino would close a room whenever it had not been visited on the last 15 clicks. A colored progress bar above every door signalled the availability level of the corresponding room.

Our environmental manipulation was varied as a *within-*

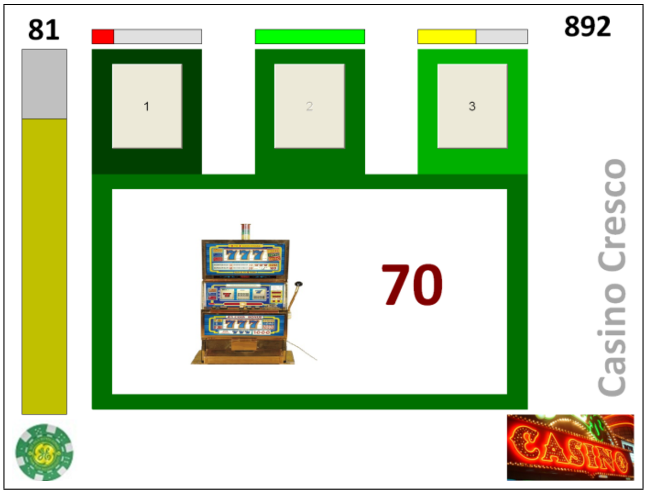

Figure 1: View of a casino in the *threatened* condition. A click on room 2's slot machine just yielded a payoff of 70 points. Clicking on doors 1 or 3 (in the upper part of the screen) would allow switching into the corresponding rooms. The horizontal bars above the doors signal the rooms' availability. (Here, room 1 would disappear soon.) Two counters (on the top left and right) denote the current budget of jetons (81 clicks remaining) and the accumulated payoff in the current casino (892 points).

subject factor. Every participant consecutively visited three casinos (in randomized order), which were rumored to employ different policies: Casino "Cresco" was reported to increase the average payoffs of its slot machines slightly as more jetons are inserted on a single room visit (a *progressive* environment). By contrast, casino "Consumo" was reported to gradually decrease the average payoffs of its slot machines (an *exhaustive* environment). Casino "Firmitas" was reported to maintain the average payoffs at the same levels (a *stable* environment). The instructions further emphasized that the payoffs of all machines were sampled from variable but mutually independent distributions, and that all casinos would reset the payoff levels of a machine whenever rooms are switched.

As the actual payoff functions of all casinos were not varied, the rumored casino policies merely changed participants' expectations about them. To obscure this fact and convey a true benefit from successful exploration of the three rooms, the payoffs were sampled from normal distributions with different means and large overlaps in values. To replicate previous results but avoid negative payoffs, we used the distributions of Shin and Ariely's (2004) Exp. 4, but linearly transformed all values to obtain mean payoffs of 50, 55, and 60 points for the worst, medium, and best rooms, respectively. To prevent transfer effects between environments, payoff values were linearly transformed for the 2nd and 3rd casino.

Participants were randomly assigned to one availability condition and the sequence of casinos and room positions was randomized for every participant. On average, an experimental session lasted for 43 minutes (with a range from 26 to 81 minutes).

Hypotheses The basic rationale of combining a multi-arm bandit problem with switch costs and an additional manipulation of option availability is the following: All participants are confronted with a learning task that poses a dilemma between exploration and exploitation. But whereas participants with *constant* availability of options only switch rooms to explore (i.e., learn the payoff functions in the other rooms), participants in with *threatened* availability of options may also switch rooms to avoid the loss of options.

Our first prediction concerns a replication of the main effect found in previous studies (Shin & Ariely, 2004; Ejova et al., 2009):

- Participants in the *threatened* condition will switch rooms more frequently than those in the *constant* condition.

Beyond this basic effect of option availability, we suggest that both the basic need to explore alternatives and the additional reason to switch to preserve options are moderated by participants' perceived decision environment. Specifically, we hypothesize two environmental effects:

- Relative to a *stable* environment, exploration will be reduced in a *progressive* environment and increased in an *exhaustive* environment (i.e., a main effect of environment on switching).
- This effect will be stronger in the *threatened* condition than in the *constant* condition (i.e., an interaction between environment and option availability on switching).

Our main dependent variable to assess these hypotheses is the number of room switches. However, the hypothesized effects also have interesting consequences for participants' accumulation of gains and the speed with which environmental properties (i.e., the switch costs and rooms' different payoff functions) can be learned and exploited.

Under the specific constraints of the paradigm the relationship between exploration and accumulated payoffs is a complex one. To maximize profits, participants need to negotiate two types of costs: (a) exploration costs that occur by forgoing payoffs when switching to another room, and (b) opportunity costs that occur when exploiting suboptimal rooms. Thus, there are two corresponding ways of increasing gains: (a) reducing the number of costly room switches, and (b) learning the environmental payoff functions to discriminate between better and worse rooms. Our background assumption that people tend to master such challenges in an adaptive manner motivates two temporal hypotheses:

- Participants will learn to reduce their number of switches with more experience.
- Participants will learn to increase their average payoffs with more experience.

In principle, these latter two hypotheses concern both availability conditions. However, it is possible that the additional need to switch rooms to keep options viable in the *threatened* condition may affect the learning process.

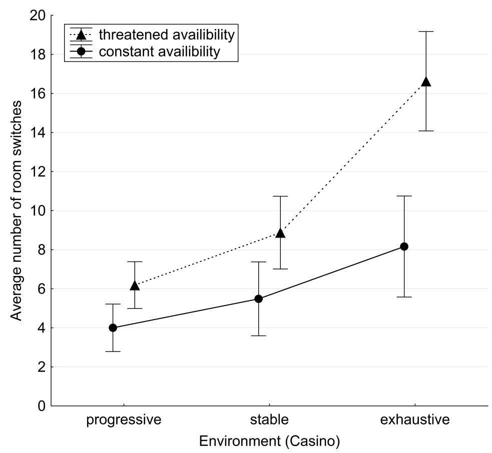

Figure 2: Mean number of room switches by environment. (Error bars in all figures indicate 95% confidence intervals.)

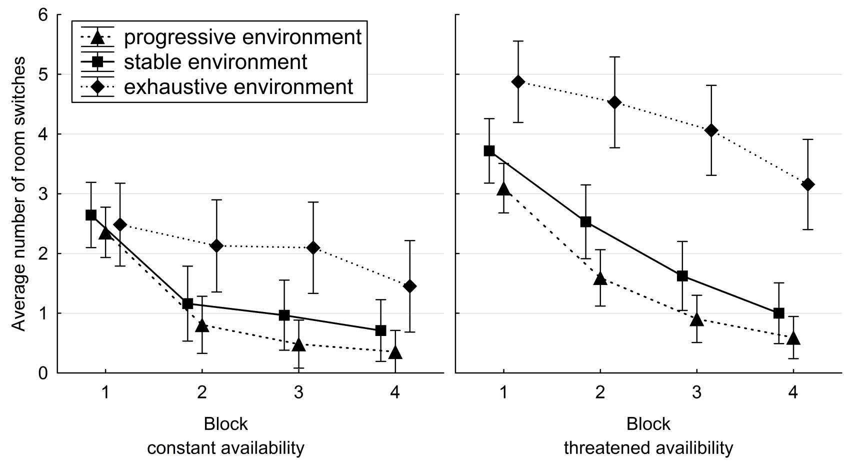

Figure 3: Mean number of room switches by block (of 25 clicks) and environments.

Results

We first consider participants' switching behavior, before investigating effects on gains and potential effects of learning.

Switching between Options

The mean number of room switches per availability condition and environment type is displayed in Figure 2. As predicted, the number of room switches was influenced by availability condition, $F_{1,61} = 20.3$, $p < .001$, $\eta^2 = .25$. In the *stable* environment, an increase in average room switches—from 5.5 in the *constant* condition to 8.9 in the *threatened* condition—replicates the basic result of previous studies. However, this basic main effect is accompanied by two environmental effects: A significant main effect of environment type, $F_{2,122} = 44.6$, $p < .001$, $\eta^2 = .42$, and a significant interaction between availability condition and environment, $F_{2,122} = 8.8$, $p < .001$, $\eta^2 = .13$. Figure 2 illustrates that the increase in room switches due to threatened options is particularly large in the *exhaustive* environment, but does not fully disappear in the *progressive* environment. When switching behavior is analyzed separately without the *exhaustive* environment, both main effects of threat condition, $F_{1,61} = 10.7$, $p < .01$, $\eta^2 = .15$, and environment type, $F_{1,61} = 8.4$, $p < .01$, $\eta^2 = .12$, still remain significant, but do no longer interact, $F_{1,61} < 1$.

Did participants reduce room switches with more experience? An ANOVA that included blocks of 25 clicks as an additional within-subject factor revealed a systematic decrease in switches across blocks, $F_{3,183} = 100.9$, $p < .001$, $\eta^2 = .62$. This main effect interacted with option availability, $F_{3,183} = 4.5$, $p < .01$, $\eta^2 = .07$, and the type of environment, $F_{6,366} = 7.3$, $p < .001$, $\eta^2 = .11$, but the three-way interaction was not significant, $F_{6,366} < 1$, $p = .97$. Figure 3 shows that the mean number of switches decreased in both availability conditions and in all three environments. Given a higher initial number of switches, this reduction was steeper in the *threatened* condition. Whereas the participants in both availability conditions reached a floor between 0 and 2 switches on the 2nd or 3rd block of *stable* and *progressive* environments, they continued to explore rooms for much longer in the *exhaustive* environment.

Learning to Distinguish Better from Worse Options

From a participants' perspective, our analysis of room switches only concerns a by-product of the explicit goal of maximizing profits. With respect to gains there are two candidate measures to consider: participants' cumulative payoffs X_{total} and their mean payoff per option choice X_{mean}. Given the paradigm's budget constraint, both measures are related as follows:

$$X_{total} = (99 - N_{switches}) \cdot X_{mean}$$

i.e., the cumulative payoff X_{total} depends on the mean payoff per choice X_{mean} and is negatively related to the number of room switches $N_{switches}$. As the three rooms' expected payoffs are 50, 55, and 60 points, a switch costs 55 points on average. This magnitude and the negative relationship between X_{total} and $N_{switches}$ implies that the paradigm punishes switching so heavily that the pattern of participants' cumulative payoffs X_{total} tends to be a mirror image of their switching behavior. (As the correlation between X_{total} and $N_{switches}$ is $r = -.90$ a complete analysis of X_{total} is omitted here.)

Despite its high costs, switching could still be worthwhile if exploration enabled participants to increase X_{mean} by learning to distinguish better from worse options. If participants never explored their X_{mean} would approximate 50, 55, or 60, depending on which room they happened to select on their first choice. Any systematic increase of X_{mean} beyond the expected value of 55 would indicate successful learning; a value of 60 would indicate perfect learning.

Figure 4 shows the values of X_{mean} per environment and condition. Importantly, participants in both availability conditions succeeded in learning payoff functions beyond baseline in all three environments. A corresponding ANOVA

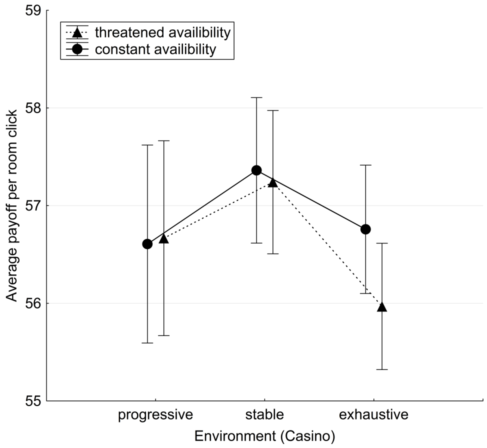

Figure 4: Mean payoffs X_{mean} by environment. (Values of 55 and 60 would indicate no and perfect learning, respectively.)

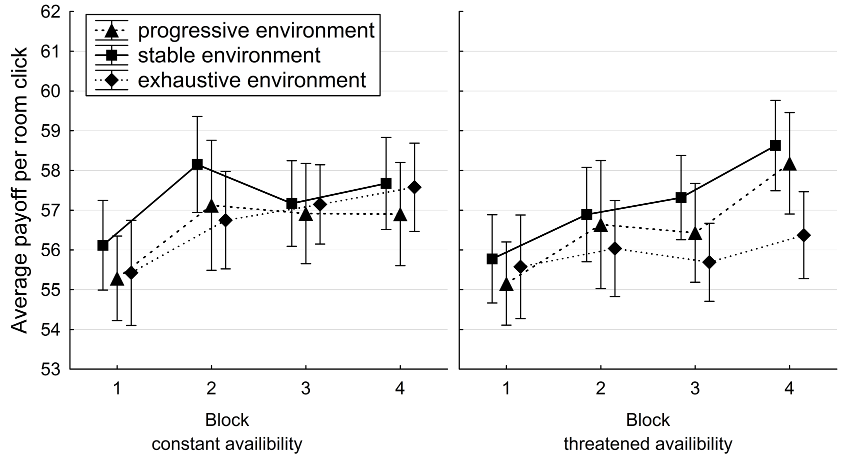

Figure 5: Mean payoffs X_{mean} by block (of 25 clicks) and environments.

yielded a small but significant effect of environment, $F_{2,122} = 3.5$, $p < .05$, $\eta^2 = .05$, but neither an effect of availability condition, $F_{1,61} < 1$, $p = .47$, nor an interaction, $F_{2,122} < 1$, $p = .47$.

To interpret this pattern and study the time course of learning, we also conducted an ANOVA that included blocks of 25 clicks as an additional within-subject factor (see Figure 5). This yielded a significant main effect of block, $F_{3,183} = 9.2$, $p < .001$, $\eta^2 = .13$, but no additional two-way (all Fs< 1) or three-way interactions ($F_{3,366} = 1.8$, $p = .10$).

When only considering the fourth block (i.e., the final 25 clicks), this revealed no influence of availability condition, $F_{1,61} < 1$, $p = .60$, but a small main effect of environment, $F_{2,122} = 3.2$, $p < .05$, $\eta^2 = .05$, and a significant interaction between environment and option availability, $F_{2,122} = 4.3$, $p < .05$, $\eta^2 = .06$. This interaction can be interpreted when comparing Block 4 of Figure 5 in both availability conditions: Whereas the average payoffs X_{mean} on the final 25 clicks did not vary between environments in the *constant* condition, they did vary in the *threatened* condition. Specifically, an average payoff of 56.2 in the *exhaustive* environment was systematically lower than the average payoffs of 58.6 and 58.2 in the *stable* and *progressive* environments.

Discussion

The observed pattern of results supports most of our hypotheses. As expected, participants switched rooms more frequently when the availability of options was threatened. This replicates the original main effect of availability condition by Shin and Ariely (2004) that was also observed by Ejova et al. (2009).

The main contribution of our study is a demonstration that this effect is modulated by people's expectations about the choice environment. Specifically, people invest additional effort into exploration in *exhaustive* environments and tend to indulge in more exploitation in *progressive* environments. As predicted, the push towards increased exploration when options are threatened is emphasized in *exhaustive* environments.

Our participants' balance between exploration and exploitation also showed clear evidence of learning. With more experience in the task environment, they rapidly reduced the number of costly switches. At the same time, participants learned to discriminate between better and worse options and used this insight to increase their average payoffs over time. Overall, this pattern of results suggests that people adapt to both an environmental manipulation and to a threat of option availability in a sensible and highly sensitive fashion.

Nonetheless, two possible challenges to our participants' ecological rationality remain: First, the possibility of losing options in the *threatened* condition led to an increased number of switches even in our *progressive* environment. This indicates that—in line with Shin and Ariely's (2004) original results—the mere threat of disappearing options encouraged participants to invest additional effort into their preservation even when their loss would no longer be problematical. The robustness of this phenomenon could indicate that people did not fully trust our instructions about casino policies, or were able to detect that the actual payoffs did not change as a function of option exploitation. At the same time, it is encouraging that increasing the number of switches in the *threatened* condition only involved a small, but non-significant loss of overall profit in the *progressive* environment.

Second, the combination of *threatened* options and an *exhaustive* environment led to a much higher number of switches and substantially lower average and overall profits. But rather than demonstrating a lack of rationality, participants in deteriorating environments are fully justified in preserving threatened options. Our study highlights that people who are forced to keep options open under such conditions generally suffer from two distinct disadvantages: Not only do they have to pay high switch costs to preserve necessary options, but their reduced experience with any particular option interferes with the learning of payoff functions.

Conclusion

The original results by Shin and Ariely (2004) suggested that the human tendency of keeping threatened options viable may constitute a form of irrational behavior. Despite replicating the original effect, our analysis and findings support the opposite conclusion.

Our basic premise was the intuition that any question about an organism's rationality can only be asked and answered in reference to properties of the decision environment in which some behavior is taking place (Simon, 1956). By adopting an ecological approach, we proposed that—depending on the dynamic properties of decision environments—both preserving and abandoning options can be adaptive and ecologically rational in real-world contexts. By extending the original bandit-problem with threatened options to different types of decision environments, we suggested that people should stay in *progressive* environments, but leave and seek alternatives in *exhaustive* environments.

Overall, our ecological stance predicted and explained a rich pattern of findings as a function of environmental expectations. As a result, an apparently irrational tendency to invest effort into the preservation of options can be re-interpreted as an ecologically rational adaptation to different types of environments. However, the results reported here are silent about the underlying mechanisms that govern the preservation and abandonment of options. In addition to Shin and Ariely's (2004) proposal of a general aversion against the "loss of options," we would not want to discard the possibility that some "desire for flexibility" may boost the attractiveness of threatened options. For instance, research on learning theory suggests that the mere restriction of a stimulus may convey reinforcing qualities (e.g., Timberlake & Allison, 1974).

In our findings, a basic bias to preserve threatened options prevailed even in progressive environments. Although this contradicted our original intuitions about the people's willingness to sacrifice options under such conditions, even this result can be considered rational under an additional assumption. In real environments, organisms are rarely guaranteed that a single option will *never* get worse. Consequently, a basic desire for keeping some options open may not be an irrational bias, but rather a sensible adaptation to a fundamentally uncertain world.

References

Banks, J., Olson, M., & Porter, D. (1997). An experimental analysis of the bandit problem. *Economic Theory*, *10*(1), 55–77.

Ejova, A., Navarro, D. J., & Perfors, A. F. (2009). When to walk away: The effect of variability on keeping options viable. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Meeting of the Cognitive Science Society* (pp. 1258–1263). Austin, TX: Cognitive Science Society.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431–440.

Lee, M. D., Zhang, S., Munro, M., & Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research*, *12*(2), 164–174.

Payne, S., Duggan, G., & Neth, H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology: General*, *136*(3), 370–380.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*, 643–675.

Shin, J., & Ariely, D. (2004). Keeping doors open: The effect of unavailability on incentives to keep options viable. *Management Science*, 575–586.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.

Sims, C. R., Neth, H., Jacobs, R. A., & Gray, W. D. (2013). Melioration as rational choice: Sequential decision making in uncertain environments. *Psychological Review*, *120*(1), 139–154.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.

Steyvers, M., Lee, M., & Wagenmakers, E. (2009). A Bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, *53*(3), 168–179.

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

Timberlake, W., & Allison, J. (1974). Response deprivation: An empirical approach to instrumental performance. *Psychological Review*, *81*(2), 146–164.

Todd, P. M., Gigerenzer, G., & the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press.