To act fast or to bide time? Adaptive exploration under competitive pressure

Douglas B. Markant

Department of Psychological Science, University of North Carolina at Charlotte

Nathaniel D. Phillips

Roche, Basel

Yaakov Kareev

Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem

Judith Avrahami

Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem

Ralph Hertwig

Center for Adaptive Rationality, Max Planck Institute for Human Development

Author Note

Douglas Markant, Department of Psychological Science, University of North Carolina at Charlotte. Correspondence concerning this article should be addressed to Douglas Markant, Department of Psychological Science, Colvard South Building, UNC Charlotte, 9201 University City Blvd., Charlotte, NC 28223. E-mail: dmarkant@uncc.edu

Word count: 14,891

Abstract

Competitive pressure affects a wide spectrum of decisions under uncertainty. It forces the individual to balance the value of gathering more information about the quality of potential choice alternatives against the risk that competitors will act first and claim the best options. Although this tradeoff between competition and exploration has long been recognized, little is known about how people adapt their exploration of uncertain options when facing competitive pressure. We examined how competitive pressure affects exploration in the *rivals-in-the-dark* game. Two players simultaneously learn about a set of choice options and compete to claim the best one. Across three studies, we show that people adapt their exploration in response to the structure of the choice environment (including the option set size and the relative number of gains and losses) and in response to repeated competition with the same opponent. Furthermore, we present a model-based analysis showing that their behavior is best described by a compensatory strategy under which the value of further exploration is weighed against the cost of being beaten to the punch by an opponent. The results point to a process of local adaptation whereby people learn to "act fast" based on their experience in a novel competitive environment.

Keywords: decisions from experience, competition, uncertainty, exploration

To act fast or to bide time? Adaptive exploration under competitive pressure

People regularly compete for resources despite being uncertain about their true value. Competitive pressure affects a wide range of decisions under uncertainty, including selecting promising investment opportunities, hiring attractive job candidates, and purchasing a home in a desirable neighborhood. In each of these examples, the individual is initially uncertain about the relative values of available options. When left alone, she may prefer to bide her time, continuing to gather information about each option, until she feels ready to make a final choice between them. Competition, however, rarely affords the luxury of such well-informed decisions. Biding time carries the risk that competitors will act first and claim the best options for themselves. People must therefore balance the value of reducing uncertainty about options against the potential cost of losing out to others in the meantime.

Take, for instance, the goal of reserving a hotel room for a weekend trip. Prospective vacationers have many opportunities to learn about the quality of potential choices. For each option, they could read about amenities on a hotel's website, look at pictures of rooms or the surrounding neighborhood, read reviews from previous customers, and so on. The extent to which people engage in such exploration, however, likely depends on whether other people are competing for the same resource. Should the trip fall during a holiday season when many other people are searching with the same goal in mind, too much time spent exploring could mean that the best options are already gone by the time a decision is reached. The power of such competitive pressure to shape decisions is seen in marketing that highlights demand for limited resources (e.g., encouraging consumers to "act fast" as supplies are "flying off the shelves").

Research in economics and organizational behavior has long recognized that the benefits of acquiring more information or experience can be counteracted by the costs of competition (Dickson, 1992; Stigler, 1961). Making fast decisions can be especially crucial for firms competing in uncertain or volatile environments (Eisenhardt, 1989). At the level of the individual, it is less clear how competitive pressure affects exploration during decision making. There is, however, strong evidence that people adapt how much they explore based on other cost-benefit tradeoffs unrelated to competition (Ratchford, 1982). People collect less information when such exploration involves costs, including monetary penalties (Busemeyer & Rapoport, 1988; Rapoport & Tversky, 1970), opportunity costs (Payne, Bettman, & Luce, 1996; Rieskamp & Hoffrage, 2008), or high degrees of effort (Fu & Gray, 2006). Conversely, people explore more when larger rewards are at stake (Hau, Pleskac, Kiefer, & Hertwig, 2008), when they experience greater variability in outcomes (Lejarraga, Hertwig, & Gonzalez, 2012), or when the set of available options is larger (Frey, Mata, & Hertwig, 2015; Hills, Noguchi, & Gibbert, 2013).

Does this adaptive exploration hold under the threat of competition? Or does competitive pressure always cause people to “act fast,” regardless of their uncertainty about the options they choose? In the present article we investigate whether people weigh the costs of losing out to competitors against the situation-specific benefits of gaining more experience. Such a compensatory cost-benefit analysis is central to the expected-utility framework assumed by game theory (Luce & Raiffa, 1957; von Neumann & Morgenstern, 1944). As a decision making strategy, it would predict that people weigh many factors when deciding how to explore, including how much they stand to learn from searching further and what might happen if someone else acts first. On the other hand, competition may evoke non-compensatory strategies that are well-adapted to social contexts even though they ignore some features of the environment (Bröder, 2000; Hertwig & Hoffrage, 2013; Rieskamp & Hoffrage, 2008). For instance, striving to always act before anyone else may be a fruitful strategy in a competitive world, even if it involves sometimes betting on options that turn out to be poor. Distinguishing these decision making strategies is important to understanding how people compete in uncertain environments, as well as the factors that push them toward fast action.

Rivals in the Dark: Balancing Exploration and Exploitation under Competitive Pressure

We examined how people adapt their exploration in response to competitive pressure using a variant of the *rivals-in-the-dark* game introduced by Phillips, Hertwig, Kareev, and Avrahami (2014). The game embeds the *sampling paradigm*, a tool for studying solitary, experienced-based choices (Hertwig, Barron, Weber, & Erev, 2004), into a strategic context. In the rivals-in-the-dark game, two players compete for the same set of choice options. Each option is a gamble that probabilistically generates a set of different outcomes (i.e., numerical values), and players can learn about options through repeated, non-consequential sampling (see Figure 1A). Players are instructed to sample until they are ready to choose one of the options for a monetary reward based on the chosen option's expected value (EV). The game therefore separates an initial phase of exploration from a final exploitative choice.

Previous studies of solitary behavior in the sampling paradigm have typically observed median sample sizes of about 16 total draws for problems with two options (for a review, see Wulff, Mergenthaler-Canseco, & Hertwig, 2018). In comparison, players in the rivals-in-the-dark game drastically curtailed their exploration in the face of competitive pressure (Phillips et al., 2014), most often making only a single draw (the minimum sample size permitted) before making a choice.

How does competition bring about this dramatic shift in behavior? The minimal exploration reported by Phillips et al. (2014) may indicate a non-compensatory response to competitive pressure: betting on the value of choosing first while ignoring the downsides of choosing based on little information. Such a strategy may be an effective response in the absence of knowledge about how opponents will behave. That is, even very small samples can provide an (modestly) informative cue to the overall value of an option (Hertwig & Pleskac, 2010), as was the case in the choice environments of Phillips et al. and the present studies. Players can gain an edge—however slight—over their opponent by prioritizing fast decisions at the expense of reducing uncertainty. Accordingly, participants in Phillips et al. (2014) who made the first choice were more likely to obtain the option with the higher EV than their opponents, even though they frequently chose on the basis of just a single observation. In analogy to other examples of fast-and-frugal decision strategies that curtail information search by exploiting environmental structure (Gigerenzer, Hertwig, & Pachur, 2011; Todd & Gigerenzer, 2012), acting on minimal information may be advantageous in competitive settings even when uncertainty about options' quality is high.

Minimal exploration does not, however, uniquely identify a non-compensatory reaction to competition. A compensatory, cost-benefit account would suggest that minimal exploration stems from how participants weighed the perceived competitive pressure against the benefits of additional experience in the environment in question. In real-world domains, the intensity of competitive pressure can vary considerably as a function of resource type, social structures (e.g., dominance hierarchies), or even time (e.g., seasonal demand). Although prioritizing fast decisions increases the chance of choosing before competitors, such a strategy may forego rewards when competitive pressure is actually low (e.g., when there are few competitors relative to the number of available options; see Phillips et al., 2014) or when small samples are misleading indicators of options' long-term values. Similarly, the value of exploration is context-sensitive in that it is informed by previous experiences, such as extremely negative outcomes that suggest an option is not worth exploring further. In light of these variations in environmental structure (e.g., highly skewed vs. normally distributed outcomes), individual experiences, and degrees of competitive pressure, people may benefit from conditioning their exploration on the local properties of the choice environment. If so, one should expect that people go beyond a one-size-fits-all response to competitive pressure by adapting how much they explore in light of these circumstances.

Our goal was to examine this process of local adaptation when exploring under competitive pressure. Our studies were designed to address three key questions. First, we investigated whether exploration was affected by the structure of the competitive environment, including the degree of competitive pressure (i.e., the ratio of available options to number of competitors) and the distribution of options' values (i.e., the relative number of gains and losses with positive and negative EVs, respectively). Second, we examined how the kind of feedback received by players influenced their willingness to explore. In particular, we tested how exploration changed when players only received social feedback (i.e., information about which player was the first to choose) or a combination of social and payoff feedback (i.e., the payoff from the option obtained at the end of each game). Third, we examined whether exploration changed across repeated interactions with the same competitor. Across multiple studies, we find that individuals under the threat of competition consistently draw small samples and commit to choices despite high uncertainty about options' quality, indicating that the results of Phillips et al. (2014) generalize to a novel choice environment. However, we also find that exploration is influenced by several properties of the competitive environment and that it changes over the course of repeated play, suggesting that people use a process of local adaptation whereby they learn to act fast in response to experiencing competition.

The organization of the article is as follows. We first describe a novel choice environment that was the basis for our task and establish the relationship between exploration and expected performance (i.e., the ability to choose the option with the highest EV). We then describe three studies testing how people adapt their exploration under different competitive conditions. Finally, we present a model-based analysis aimed at understanding how participants make round-by-round decisions in the rivals-in-the-dark game. The model results provide further insight into our empirical findings by testing whether participants weigh the costs of competition against the benefits of further exploration in a context-sensitive, compensatory manner.

Choice Environment

The value of exploring an option through repeated sampling depends on the structure of a given choice environment, including the distribution of potential outcomes and their respective probabilities. At one extreme, if the first draw from an option is a perfectly valid cue of its value (e.g., as with a "sure thing" that always generates the same outcome), nothing is gained from sampling it more than once. In contrast, research involving the sampling paradigm often employs options with relatively consequential *rare events*, that is, infrequent outcomes that are unlikely to occur in small samples but have a large impact on an option's overall quality. In these environments, small samples are likely to be insufficient to accurately assess an option's EV.

How do extreme rare events influence exploration? Existing theories of exploratory choice point to two kinds of potential mechanisms (Cohen, McClure, & Yu, 2007; Wilson, Geana, White, Ludvig, & Cohen, 2014). The first view is that exploration functions much like exploitative choice in that attractive options are explored more frequently than unattractive options (Gonzalez & Dutt, 2011; Lejarraga & Hertwig, 2016). Rare outcomes have no influence on exploration until they are experienced and change the perceived value of an option. The second view is that exploration is driven by beliefs about the environment, and in particular, predictions about how much information will be gained from different exploratory actions (Markant, Settles, & Gureckis, 2016; Nelson, 2005). The role of this latter belief-driven process in previous studies of the sampling paradigm is unclear. Participants are typically not informed about how options are generated, including the potential for extreme rare events (although some participants may discern this structure across multiple problems and subsequently prolong search in order to discover unexperienced rare outcomes, see Mehlhorn, Ben-Asher, Dutt, & Gonzalez, 2014).

We aimed to test whether people trade off the costs of competition against the predicted benefits of exploration. To that end, we designed a choice environment in which the value of learning about rare outcomes through repeated sampling was transparent. In our variant of the rivals-in-the-dark game participants were informed about the probabilities and possible ranges of outcomes. Each option was associated with two outcomes: a *common* outcome that occurred with probability $p = .8$ and a *rare* outcome that occurred with probability $1 - p = .2$. The common outcome was a single number randomly sampled from a uniform distribution of discrete values in the range [-20, 20], whereas the rare outcome was sampled from a uniform distribution of discrete values in the range [-200, 200] (see Figure 1B for an illustrative problem).

The option environment was thus characterized by potentially high-magnitude but infrequent outcomes that could have a large impact on the quality of an option. Given a choice between two options H and L with higher and lower EVs, respectively, players' chances of choosing the H option increased substantially if they sampled enough to experience the rare outcome for at least one of the two options. To illustrate the value of experiencing a rare outcome in this environment, we generated a set of 10,000 two-option problems and assessed how experiencing different subsets of outcomes for options H and L affected the probability of choosing the option with higher EV, $p(H)$ (see Appendix A for details). For a given problem and set of observations, the predicted choice was derived from a Bayesian ideal observer that chooses the option with the higher predicted EV based on observed outcomes. The resulting $p(H)$ is shown for all subsets of experienced outcomes in Table 1. When no rare outcome values are experienced, the proportion of H choices based on observing a single common outcome is .59, and increases slightly to .62 when common outcomes from both options are experienced. In contrast, $p(H)$ is .81 when only a single rare outcome is observed and increases further as common outcomes are experienced for both options. Finally, observing the rare outcomes from both options leads to $p(H)$ approaching 1, even when neither common outcome has been experienced. Thus, although experiencing one or both common outcomes across the two options results in better-than-chance selection of H, players' ability to choose H rises substantially if they sample long enough to observe at least one rare outcome.

We next examined how choice performance depends on the amount of exploration in terms of total sample size. The probability of observing at least one rare outcome is described by the cumulative geometric distribution in Figure 1C (solid line). The expected probability of selecting H was found by simulating sets of observations of a given sample size and using the same choice procedure as above (Appendix A). The resulting mean $p(H)$ for the ideal observer is indicated by the dashed line in Figure 1C. Note that 16 draws, the median sample size observed in previous studies of the sampling paradigm, is associated with near perfect accuracy (99% chance of choosing H) in our choice environment under this model. As such, highly accurate choices could be made with a modest amount of exploration, provided, however, that the individual experienced one or more of the rare outcomes.

Study 1: Distinguishing the Threat of Competition from Opportunity Costs

Our first goal was to evaluate the effects of competitive pressure in the choice environment described above. Relative to behavior in solitary implementations of the sampling paradigm (e.g., Hertwig et al., 2004), we anticipated that sample sizes would be smaller due to participants' explicit knowledge about the environment, which in most problems obviates the need to estimate outcomes' relative frequency through repeated sampling.[1] At the same time, we predicted that sample sizes may be larger than seen by Phillips et al. (2014) because the outcome distributions in the present environment are more skewed (i.e., higher variances in outcomes; see Lejarraga et al., 2012) and participants were explicitly informed about this potential for high-magnitude rare outcomes.[2]

The second goal was to examine whether reduced search under competitive pressure is due to increased opportunity costs imposed by the synchronous nature of the rivals-in-the-dark game. In solitary settings people adjust how much they search based on the costs involved in obtaining information, whether in the form of high degrees of effort (Gray & Fu, 2004), monetary penalties (Busemeyer & Rapoport, 1988; Rapoport & Tversky, 1970), or opportunity costs (Payne

[1] In most problems the rare outcome falls outside the range of common outcomes ($[-20, 20]$), and can therefore be identified as occurring with $p = .2$ after a single observation. However, for some problems the rare outcome happens to fall within the same range as the common outcome. In those cases, repeated sampling would still be necessary to estimate the relative frequency of the two outcomes.

[2] In Phillips et al. (2014) each option was associated with a positive outcome and a negative outcome. Across the full set of gambles, these outcomes fell within smaller ranges than in our environment (positive: [25, 55], negative: [-26, -11]) whereas the probability of the positive outcome ranged from .22 to .5 ($M = .35$). As a result, lower-probability outcomes tended to be both less extreme and more likely to occur in small samples as compared to the present environment.

et al., 1996; Rieskamp & Hoffrage, 2008). Competitors in the rivals-in-the-dark game are subject to delays from waiting for opponents to make decisions. These delays represent additional opportunity costs that may decrease search effort but are not specific to the competitive nature of the interaction. To test this alternative explanation, we compared a *Competitive* condition to an *Independent* condition. In the former, players' choices were dependent on the actions of their opponents; in the latter, participants observed the final choices of a partner but were free to search and choose options independently. Opportunity costs and social feedback (cues indicating partners' choices) were matched across conditions, allowing us to isolate the effect of competitive pressure on search effort and final choices.

Finally, we employed a repeated-play design to explore the dynamics of competitive search across multiple games with the same opponent. Will competitors adopt an unbending sampling strategy from the outset of the experiment, that is, prioritizing fast decisions by always stopping after a single observation? Or, will they adapt their search effort in response to competition, for instance, after they experience an opponent choosing first? Given the consequential nature of social feedback experienced during competitive play, we expected that changes in sample size would not be observed when the same feedback did not constrain participants' decisions (Independent condition).

Participants and Materials

We recruited 212 participants through Amazon Mechanical Turk (http://www.mturk.com) using the *psiTurk* software package (http://www.psiturk.org; Gureckis et al., 2015). Forty-four participants (20%) failed to complete the game because a member of their group left early, leaving a total of 168 complete experimental runs (61 female, 54 male, 1 other, 52 no response;

$M_{age} = 35.6$, $SD = 11.5$, 52 no response). Participants received a base payment of $0.50 for their participation as well as a bonus of up to $3 depending on their performance. Participants were randomly assigned to either the Independent condition or Competitive condition (both $N = 84$ or 42 pairs in each condition).

We randomly generated 20 problem sets comprised of 8 problems each using the choice environment described above. Option sets were resampled if the difference between the EV for the *H* and *L* options was less than 25 points. In addition, option sets were resampled if the summed EV across all *L* options was less than -100 or the summed EV of all *H* options was greater than 200, ensuring that each participant's total number of points at the end of the experiment would lie within that range.

Procedure

Practice. Participants were informed about the probabilities and ranges of the common and rare outcomes, and were instructed that the goal of the game was to claim the option with the higher average outcome value. Prior to playing, they completed four trials of non-consequential, solitary sampling with individual options. In each trial, a single option appeared and participants were instructed to sample 25 times and observe the resulting outcomes. They were then asked to report the two outcomes observed during sampling and to estimate the average value of the observed outcomes. All participants experienced the same four practice options, including options where both outcomes came from the same domain (e.g., both the common and rare outcomes were positive) and options where they came from different domains (e.g., a common negative outcome but rare positive outcome with high value). This practice ensured that participants were familiar with the structure of individual options, including the relative frequency and range of each outcome type, as well as the EV criterion we encouraged them to maximize.

Group formation and coordination. Upon completion of the practice trials, participants were presented with a list of open groups and instructed to join one group. After joining an open group, they waited for another person to join, at which point both confirmed that they were ready to play. If a participant waited for more than 15 minutes without a second person joining their group, the experiment aborted and they were paid for partial participation. Gameplay was coordinated such that the game advanced only when decisions were received from each participant and broadcast to the other. This helped to ensure that both participants were present and attentive throughout the experiment. For example, participants continued to the next sampling round only when they acknowledged their opponent's decision to either stop or continue sampling. If either participant closed the experiment or was idle for more than 4 minutes at any point after joining a group, the experiment ended and both participants were paid for partial participation. Only data from the 84 pairs that completed the full set of eight games are analyzed.

Gameplay. Participants began with an endowment of \$1.00 and were instructed that their payoff from each game (the EV of the option they selected) would be added or subtracted to determine their final bonus, with each point corresponding to \$.01. One of the twenty problem sets was randomly selected for each pair of participants and the eight games were played in a random order. On each round of the game, a participant clicked on one of the two options (displayed as two urns filled with coins) and observed a randomly generated outcome (a coin labeled with a number of points between -200 and 200, randomly sampled according to the underlying distribution for that option). The outcome remained visible until the participant indicated whether they wanted to "continue learning" or "stop and choose" by clicking one of two buttons at the bottom of the display (see Figure 1A). If both participants decided to continue sampling the game proceeded to the next round. When participants decided to stop and choose an option, they then clicked on one of the two options to claim it (subject to the condition-specific procedures below). No feedback about the payoff from the chosen option was provided during the games. At the end of the game play, participants were shown the true value of the options they chose and their total bonus.

Independent condition. In the Independent condition, participants' final choices were made known to their partners, but a partner's choice had no other consequences on the participant's ability to sample or choose options. When participants decided to stop, their selections were visible to their partners in the form of an icon that appeared below the chosen option. However, partners were able to continue sampling for as many turns as they desired, and when they stopped could select either of the two options. Both participants in a pair were still required to acknowledge the completion of every round, ensuring that the participant who stopped first continued to pay attention to her partner's behavior for the remainder of the game and that the opportunity costs were matched with those of the Competitive condition.

Competitive condition. In the competitive condition, a participant's decision to stop and choose an option removed it from the set of options available to his or her opponent. When such a choice occurred, the option faded out on the display and the opponent was required to immediately select the remaining option. If both participants decided to stop on the same round, a random choice order was generated to determine which participant went first. That participant was awarded the value of the chosen urn whereas the other participant was awarded the value of the remaining urn. All other aspects of the game were the same as in the Independent condition.

Results

Exploration. Mean sample size across eight games is shown in Figure 2A. We used mixed effects negative binomial regression to evaluate the effects of condition (Independent vs. Competitive) and trial number (games 1–8) on sample size (using the *lme4 R* package), with a random effect to model variability across pairs. Table 2 reports the parameter estimates, confidence intervals, and inferential statistics for the resulting model (with the Independent group as the reference condition). Each estimated fixed effect indicates the change in log sample size associated with a unit change in the predictor. There was an effect of condition indicating smaller sample sizes in the Competitive condition ($M = 2.1$, $SD = 1.8$) compared to the Independent condition ($M = 5.1$, $SD = 4.7$). In addition, there was a positive effect of trial number on sample size, indicating an increase in sample size over games in the baseline Independent condition. Finally, there was an interaction between the Competitive condition and trial number. A post-hoc contrast indicated, in contrast to the Independent condition, there was an overall decrease in sample size over games in the Competitive condition ($\beta = -.11$ $[-.17, -.06]$, $p < .001$).

The decrease in sample size over games within the Competitive condition was evident at the level of individual pairs, with 27 pairs (64%) showing a decrease in mean sample size from the first half to second half of the game rounds. Of the remaining pairs, 6 (14%) showed no change, and 9 (21%) showed an increase in mean sample size. In contrast, pairs in the Independent condition showed the opposite pattern of change from the first to second half, with 15 pairs (36%) showing a decrease, 1 (2%) showing no change, and 26 (62%) showing an increase in sample size.

The difference in sample size between conditions affected the likelihood of experiencing rare outcomes, as anticipated by our description of the choice environment. In the Independent condition, participants did not observe either rare outcome in 40% of all games, whereas participants in the Competitive condition (combined across all participants regardless of whether they decided to stop or not) did not observe either rare outcome in 65% of all games ($\chi^2 = 84.3$, $p < .001$). Note that it was not the case that competitors who decided to stop were simply more likely to have observed a rare outcome. Focusing on those games in which one participant stopped before the other (non-ties), in 55% of them the first stopper (i.e., the participant from each pair who was the first to terminate sampling) did not experience a rare outcome. Among tied games, at least one participant had not observed a rare outcome in 81% of games. Combining these cases leads to 63% of games in which at least one participant in a group decided to stop before observing a rare outcome.

Notably, although sample size was higher in the Independent condition on the first game, the mean sample size among first stoppers (players that were the first of a pair to stop exploring) in that condition was similar to that of the Competitive condition (see dashed line in Figure 2). Negative binomial regression was used to test the effect of condition on sample size in the first game, focusing only on the sample size of the first stoppers in each group. There was no effect of condition ($\beta = .03$, $z = .17$, $p = .86$). This result suggests that participants in both conditions may have began with similar strategies, but differentially adjusted their exploration across trials due to the presence or absence of competition.

Was the decline in exploration in the Competitive condition a response to losing out to opponents? We used mixed effects logistic regression to model whether sample size decreased between successive trials (binarized). The main factor of interest was whether first choosers on trial t were “slower” than their opponent on trial $t-1$ (Slower_{t-1}), that is, whether they had been the second chooser in the previous trial. In addition, trial number $(2-8)$ and the Trial $\times$

Slower_{t-1} interaction were included as predictors, since decreases between successive trials were less frequent in later trials. The results are shown in Table 3. Effect sizes are reported in terms of relative odds ratios (*OR*). As compared to cases in which the chooser was the same across trials, being beaten by the opponent on the previous trial was associated with a higher likelihood that sample sizes decreased. This effect was largest on the second trial: being beaten on the first trial was associated with an increase of a factor of $OR = 3.36\ [1.38, 8.58]$ in the relative odds of sample size decreasing on the second round. This is strong evidence that being out-chosen is a critical factor in reducing search in competitive environments.

Final choices. We next evaluated whether the two conditions differed in their selection of the *H* option. Figure 2B shows the proportion of *H* (rank=1) and *L* (rank=2) choices among first and second choosers across all games, for games in which neither rare outcome was experienced, and for games in which at least one rare outcome was experienced. In the Independent condition, both first and second choosers were more likely than not to choose the *H* option and benefited from observing at least one rare outcome. In the Competitive condition, a similar advantage was only seen for first choosers.

We used mixed effects logistic regression to model the effects of condition (Independent vs. Competitive), trial (1–8), and experiencing at least one rare outcome on the likelihood of choosing *H* (Table 3). For the Competitive condition, we included only those participants who made the first choice (either because they stopped before their opponent or it was a tie and they were randomly selected to choose first). There was no effect of trial number in the Independent condition. Although the overall effect of condition was not significant (Independent: $M = .78$, $SD = .17$; Competitive: $M = .68$, $SD = .27$), there was a condition $\times$ trial interaction such that first choosers in the Competitive condition were less likely to choose *H* over the course of the experiment ($OR = .84$ [.74, .97]). Finally, as expected from our analysis of the choice environment (Figure 1), the likelihood of choosing H increased if the chooser experienced at least one rare outcome ($OR = 2.91$ [2.12, 4.03]).

Discussion

Compared with the non-competitive social context of the Independent condition, competitive pressure reduced exploration, consistent with Phillips et al.'s (2014) observations. The higher search effort observed in the Independent condition suggests that small sample sizes in the Competition condition did not result from higher opportunity costs (e.g., delays due to waiting for an opponent to make decisions) or social information alone (i.e., notification of the other participant's final choice), elements that were matched across conditions.

Small sample sizes in the Competitive condition predictably led participants to fail to experience any rare outcomes before a majority of stopping decisions. This significantly lowered their ability to choose the H option (Figure 2B). Of course, this does not imply that competitors failed to understand the potential impact of rare outcomes or acted unreasonably by stopping before they had identified any rare outcomes. Given that knowledge of the common outcomes alone can lead to better than chance selection of the H option, participants may have prioritized fast decisions based on that partial knowledge rather than risk losing the ability to make the first choice.

Finally, sample sizes changed over the course of repeated games in different directions for the two conditions. In the Competitive condition sample sizes declined, consistent with a dynamic process whereby competitors adjusted how much they sampled in response to experiencing competitive pressure. In contrast, there was a small increase in sample size in the Independent condition, indicating that the decline seen in the Competitive condition was not simply due to experience with the game or choice environment. Interestingly, the distributions of sample sizes among first stoppers in the first game were not significantly different between the two conditions, suggesting that participants in both conditions began the game with similar strategies but then diverged in how they responded to social feedback about their partners' choices.

In sum, the results of the first experiment suggest that the response to competitive pressure changes as a result of direct experience with competition. In addition, they provide clear evidence that competitive pressure leads to restricted exploration despite high uncertainty about the quality of the available options. In contrast to the procedure of Phillips et al., (2014), participants knew there were rare outcomes to be discovered in every problem. Nevertheless, they frequently stopped exploring before learning about those outcomes. Does this reflect a non-compensatory strategy of acting before competitors at the expense of reducing uncertainty? In the next two studies we examine whether this response is invariant to manipulations of the situation-specific tradeoff between competitive pressure and exploration.

Study 2: The Impact of Set Size and Payoff Feedback

Two properties of the first study may have amplified the impact of competition on exploration. First, the high ratio of players to the number of options (2:2) likely contributed to a keen sense of competition for the best option, particularly since gains and losses were equally likely. In Study 2, we examined this further by manipulating the number of choice options (2 vs. 4), with the prediction that competitors would be more willing to bide their time when more options were available, that is, when they face a "buyers' market." This prediction is supported by simulations conducted by Phillips et al. (2014, see their Figure 6) showing that players benefit from planning to collect larger samples when there are more options than competitors.

Second, participants in Experiment 1 did not receive immediate “on-line” feedback about the outcomes of their final choices (i.e., the payoffs applied to their bonus at the end of the game). Without this feedback, they could not evaluate the consequences of their choices with respect to their level of uncertainty (e.g., after claiming an option for which the rare outcome was unknown). At the same time, competitors did receive immediate on-line social feedback about whether they were able to choose first or were beaten to the punch by their opponents. In the absence of payoff feedback, social feedback may have encouraged participants to prioritize choosing first regardless of their uncertainty about the options. We tested this hypothesis in Study 2 by manipulating the type of on-line feedback participants received over the course of the game. The *No-feedback* condition was identical to the Competitive condition in Experiment 1. In the *Partial-feedback* condition, participants learned on-line (after each choice) the payoff of the option they chose but did not see the payoffs of options chosen by their opponents. If feedback about their choices permits individuals to learn the value of exploring until rare outcomes are experienced, larger sample sizes would be expected relative to the No-feedback condition. Finally, in the *Full-feedback* condition, players were given feedback about the value of both options chosen by either player. Since full feedback allows participants to directly assess how their ability to choose the *H* option depends on whether they experienced rare outcomes, we anticipated that sample sizes in the Full-feedback condition would increase to an equal or greater extent than in the Partial-feedback condition.

Participants and Materials

We recruited 618 participants through Amazon Mechanical Turk. One-hundred twenty-eight (21%) failed to complete the task because a member of their group left early or was idle for more than four minutes, leaving a total of 490 complete experimental runs (190 female, 161 male, 1 other, 138 no response; $M_{age} = 34.5$, $SD = 10.9$, 137 no response). Participants received a base payment of \$.50 for their participation, as well as a bonus of up to \$3 depending on their performance. Participants were assigned to one of four conditions based on crossing the number of options and feedback condition (2-options, No-Feedback: $N = 80$; 2-options, Partial-feedback: $N = 70$; 2-options, Full-feedback: $N = 86$; 4-options, No-Feedback: $N = 86$; 4-options, Partial-feedback: $N = 78$; 4-options, Full-feedback: $N = 90$).

Procedure

The option sets from Study 1 were used for the 2-option conditions. For the 4-option conditions, new option sets were generated with the same general procedure as detailed in Study 1, with the additional constraint that each problem included two losses (one with an EV less than -20 and a second with an EV in the range $[-20, -1]$) and two gains (one with an EV greater than 20 and a second with an EV in the range $[1, 20]$). Participants were not informed about this distribution of option EVs. All other aspects of the instructions and practice trials were the same as in Study 1.

All participants completed eight trials. Gameplay in the No-feedback conditions was identical to that of the Competitive condition in Experiment 1. In the Partial-feedback condition, participants observed the EV of the option they chose at the end of each game. In the Full-feedback condition, the EV of both chosen options was displayed to both players upon completion of each game.

Results

Exploration. Mean sample size across eight trials is shown in Figure 3A for each condition. Negative binomial regression was used to evaluate the effects of feedback condition (No-, Partial-, and Full-feedback), number of options (2 vs. 4), and trial number (1–8) on sample size (Table 2). Sample size was higher in the 4-option groups (No-feedback: $M = 2.79$, $SD = 2.2$; Partial-feedback: $M = 2.63$, $SD = 1.59$; Full-feedback: $M = 2.43$, $SD = 1.98$) compared to the 2-option groups (No-feedback: $M = 1.67$, $SD = 0.94$; Partial-feedback: $M = 1.81$, $SD = 1.22$; Full-feedback: $M = 1.53$, $SD = 0.86$), showing that participants explored more when a larger number of options were available. In addition, there was a negative effect of trial such that sample size declined over the course of the game. There was no effect of feedback condition or interaction between feedback condition and number of options. Across all 2-option games, the first chooser stopped before observing a rare outcome in 65% of games. In 4-option games, the first chooser stopped before experiencing a rare outcome in 49% of games ($\chi^2 = 52.9$, $p < .001$). Thus, although 4-option participants tended to sample longer than 2-option participants, they nevertheless frequently stopped to claim an option without experiencing any rare outcomes.

Logistic regression was used to test whether being beaten by the opponent on the previous round was associated with decreases in sample size (Table 3). Decreases were less likely in later trials ($OR = 0.87, [0.79, 0.95]$) and were more likely in the 4-option condition than the 2-option condition ($OR = 1.77,\ [1.43, 2.20]$). As in Study 1, being beaten by the opponent on the previous trial increased the likelihood that sample sizes decreased. For instance, being the second chooser in trial one increased the odds of a lower sample size in trial two by factor of

$OR = 2.27$ $[1.58, 3.32]$.

Final choices. The proportion of choices by rank (with 1 indicating the option with the highest value, H) is shown in Figure 3B for both the first chooser and second choosers, collapsed across feedback conditions. The probability of the first chooser selecting the H option was modeled using mixed effects logistic regression, with choices in the 2-option and 4-option conditions analyzed separately (Table 3). As in Study 1, observing at least one rare outcome increased the probability of choosing H among both 2-option ($OR = 2.82,\ [2.07, 3.89]$) and 4-option ($OR = 2.14,\ [1.64, 2.80]$) participants. There were no effects of feedback condition, trial, or feedback $\times$ trial interaction on choice proportions in either condition.

Discussion

As in Study 1, sample sizes declined over games in both 2-option and 4-option games. Contrary to our predictions, we did not find any effect of choice feedback on exploration or the proportion of H choices. Whereas we expected that choice feedback about one or both options would encourage participants to explore until they experienced rare outcomes, sample sizes were equivalent across feedback conditions (and in fact, were smallest in the Full-feedback group). Thus, we found no evidence that providing choice feedback counteracts the downward pressure of competition on exploration.

In contrast to the null effect of feedback, exploration was strongly affected by the number of choice options. In individual settings, sample size is roughly linearly related to the number of available options, with people exploring more as the option set size increases (Frey et al., 2015; Hills et al., 2013). Although sample sizes in the present study were small relative to those cases, we nonetheless found that competitors explored more when the size of the option set was doubled. This finding agrees with the simulation results of Phillips et al. (2014) showing that players benefit from planning to collect larger samples when a larger number of options are available, given a fixed number of competitors.

What explains the willingness of 4-option participants to explore longer than participants in 2-option games? One possibility is that, as we suggested was the case in Study 1, participants began the game with the same strategy they might have employed as a solitary player. That is, participants in the 4-option conditions may have initially explored to a larger extent (consistent with the effect of option set size in solitary settings), but then decreased their search effort across trials as a result of experiencing competition.

Alternatively, increased exploration in 4-option games may have reflected a lower degree of perceived competitive pressure in that environment. There are at least two ways that participants may have arrived at such a judgment. First, they might predict a low cost of choosing second in the 4-option case, since they can still select among the remaining options. Given the distribution of option EVs (two gains and two losses), even when their opponents chose the *H* option, participants still had a shot at selecting an option with a positive EV. Participants were not informed about this distribution, but may have reasonably assumed from the instructions and practice trials that gains and losses were equally likely. Second, 4-option participants may have predicted that it would take opponents a larger number of draws to discover an attractive option, whereas for 2-option problems even sampling a single option can be decisive as to which option should be claimed. Both explanations suggest that expectations about the distribution of option EVs influence how people evaluate the degree of competitive pressure. Next, we directly test this possibility by manipulating the ratio of gains and losses across different games in the final study.

Study 3: Competing for Few or Plentiful Gains

How is exploration affected by knowledge of the distribution of option values? This study involved two within-subjects conditions that determined the ratio of options with positive EV (gains) and negative EV (losses) in each game: 1-gain/3-loss games and 3-gain/1-loss games. All participants played 4-option games under competition and were informed about the ratio of gains and losses within each game. We considered two competing predictions for how the gain-loss ratio could affect search. On the one hand, if a higher proportion of gains leads participants to believe that opponents will sample less (e.g., because it will take fewer samples to discover an attractive option), this increased competitive pressure should cause exploration in the 3-gain condition to be lower than that of the 1-gain condition. On the other hand, if participants judge the cost of "losing" to an opponent (i.e., choosing second) to be lower when most options are gains, this should decrease competitive pressure and cause exploration in the 3-gain condition to be higher than that of the 1-gain condition.

Participants and Materials

We recruited 152 participants through Amazon Mechanical Turk. Thirty-two participants (21%) failed to complete the task because a member of their group left early or was idle for more than four minutes, leaving a total of 120 complete experimental runs (53 female, 34 male, 33 no response; $M_{age} = 33.4$, $SD = 11.1$, 33 no response). Participants were paid a base payment of \$1.00 for complete participation (\$0.50 for partial participation) and a bonus of up to \$3 depending on their performance.

Procedure

All participants were assigned to competitive play in games with four options and no on-line payoff feedback. Participants were instructed that the value of each option could either be positive (a gain) or negative (a loss) and that games would vary in the ratio of the two types. During each game, the number of options from each domain was described above the displayed options (either “1 gain, 3 losses” or “3 gains, 1 loss”). Each pair experienced four games in each condition presented in random order.

Twenty new problem sets were generated with 8 problems per set. Each set contained 4 problems with 1 gain and 4 problems with 3 gains (otherwise losses). Problem sets were resampled if the summed value of all options with the lowest values fell outside the range [-250, -150] or the summed value of all best options fell outside the range [150, 250]. This constrained the range of final bonuses while ensuring that all option sets featured a wide range of outcome values. Each pair of participants was randomly assigned one of the twenty problem sets.

Results and Discussion

Exploration. Across all 1-gain games, the first chooser stopped before observing a rare outcome in 55% of games, whereas in 3-gain games the first chooser stopped before any rare outcome in 61% of games, a non-significant difference ($\chi^2 = 2.47$, $p = .12$). The results of mixed effects negative binomial regression on sample size (Table 2) indicated a significant negative effect of number of gains, with sample size lower in 3-gain games ($M = 2.28$, $SD = 1.64$) as compared to 1-gain games ($M = 2.74$, $SD = 2.24$). Thus, competitors searched more in environments with a high proportion of negative options compared to those with a low proportion. Unlike the previous studies, there was no effect of trial number on sample size (Table 2), indicating there was no evidence that exploration changed over the course of the experiment. In addition, losing out to opponents had no effect on the likelihood of sample size decreasing in the following trial (Table 3).

Final choices. The proportion of choices by option rank are shown in Figure 4 for both the 1st and 2nd chooser in each condition. As in the previous studies, mixed effects logistic regression (Table 4) revealed that the probability of obtaining the *H* option increased when a rare outcome had been observed ($OR = 2.63$ $[1.80, 3.88]$). In addition, there was an effect of condition such that a smaller proportion of *H* choices were made in 3-gain games ($OR = .65$ $[.43, .95]$), consistent with the lower sample sizes in that condition.

In sum, competitors' exploration was affected by the ratio of gains and losses, with an increase in sample size when faced with a single gain among four options as compared to a single loss. This result is consistent with the hypothesis that participants perceived lower competitive pressure in the 1-gain condition. That is, they behaved as if they expected that opponents would have to search more to find an attractive option. However, it cannot be ruled out that the difference in sample sizes was caused by a simpler, non-compensatory strategy that does not involve reasoning about or predicting competitors' behavior. In solitary conditions people explore more when they experience negative outcomes (Lejarraga & Hertwig, 2016; Lejarraga et al., 2012). Phillips et al. (2014) also observed that competitors were more likely to sample more than once when the first outcome they experienced was negative. Accordingly, participants in Study 3 may have simply been more likely to continue sampling in 1-gain games because negative outcomes were encountered more frequently. Distinguishing between these explanations requires a closer examination of participants' decisions in the context of the outcomes they experienced. We turn to this in the following model-based analysis.

Modeling Adaptive Exploration under Competition

Our results across three studies show that people adapt how they explore based on the local properties of a competitive environment, such as sampling more when more options are present or when losses are more common than gains. Exploration also changes in response to competitive experience, with sample sizes decreasing across repeated games against the same opponent. In this section we return to a central question regarding how people decide when to stop exploring, namely, whether they weigh the costs of competition against the value of exploration on a round-by-round basis.

Predicting how an opponent will act is key to determining how much to explore. Phillips et al. (2014) demonstrated through simulation that players in a choice environment similar to ours will tend to benefit from choosing first, even when relying on a single outcome. The simulation assumed that players begin the game by deciding on a desired sample size. If players knew their opponents would sample N times, they should sample one fewer times ($N-1$) in order to learn as much as possible about the options while preserving the ability to choose first. If they don't know how much opponents will search, however, players should err on the side of sampling too little rather than too much. High uncertainty about how long an opponent plans to sample tends to favor extremely limited exploration, particularly when the number of options is low relative to the number of players (Phillips et al., 2014, see pg. 115).

One limitation of the previous approach is that it does not account for the round-by-round nature of players' decisions in the task. Importantly, the value of exploring depends on both the likely actions of competitors and the outcomes that have been observed so far. For instance, in the present choice environment, if both the frequent and rare outcomes of an option have been seen (e.g., observing -100 and 10 from the same option), there is no uncertainty and nothing more to learn from sampling that option further. Even if there is uncertainty about an option's value, however, the benefits of reducing it may be moot if the opponent is likely to stop on the current round.

We conducted a model-based analysis to test whether participants' round-by-round decisions were driven by this tradeoff between competitive pressure and the value of exploration. We considered three decision models. All three models rely on Bayesian updating to represent uncertainty about available options and to predict the value of both immediates choices and further exploration. They differ in how these predictions drive decisions to stop and choose or to continue exploring. We briefly introduce each model here before describing the analysis in detail below.

- The *Constant* model simply assumes that players stop exploring with a constant probability following each draw, implying that decisions to stop are independent of the outcomes that are experienced. If, for example, players adopt a non-compensatory strategy in which they stop after a single draw regardless of its value, their behavior would be captured by the Constant model with a high stopping probability.

- The *Choice-first* model assumes that people decide whether to stop and choose based on options' predicted values, such that an option with a high predicted value is likely to be chosen immediately; otherwise sampling continues. This strategy echoes Phillips et al.'s observation that many participants stopped immediately when the first outcome they observed was positive (suggesting an attractive option), but continued exploring further when it was negative. It is also consistent with findings of larger sample sizes when negative outcomes are experienced in solitary implementations of the sampling paradigm (Lejarraga et al., 2012).

- Finally, the *Tradeoff* model assumes that players weigh the value of an immediate choice against the value of making an additional draw, explicitly accounting for the probability that the opponent will stop and claim an option first. Unlike the Constant and Choice-first models in which exploration is contingent on a previous decision to *not* stop and choose an option, under the Tradeoff strategy the player simultaneously compares the expected payoff of an immediate choice against that of exploring further given the costs imposed by competition. In solitary sequential decision making with fixed costs of collecting information, Busemeyer and Rapoport (1988) found that stopping behavior was best described by a similar *myopic* strategy, in which the expected payoff of an immediate exploitative choice is compared to that expected after one more round of exploration.[3] It is an open question, however, whether people rely on a similar strategy when the costs of exploration arise from competition.

Belief Updating

The decision models use Bayesian updating to represent uncertainty about the set of available options, $O = \{A, B, ...\}$, given the outcomes observed so far, X. The goal is to identify the state of option k, $s_k \in S$, where each state is a unique combination of a rare and frequent outcome from their respective ranges, $z^s_{common} \in \{Z_{common} : -20 \ldots 20\}$ and $z^s_{rare} \in \{Z_{rare} : -200 \ldots 200\}$. The hypothesis space S comprising possible option states is

[3]This strategy is considered "myopic" because it only evaluates what will happen one step into the future, whereas an optimal solution would consider action sequences of any length.

therefore the cartesian product $Z_{common} \times Z_{rare}$. Each option state is associated with a reward equal to the expected value, $\mu(s) = 0.8 \cdot z^s_{common} + 0.2 \cdot z^s_{rare}$.

The likelihood function defines the probability of observing an outcome z given a particular option state s:

$$p(z|s) = \begin{cases} .8, & \text{if } z = z^s_{common} \neq z^s_{rare}, \\ .2, & \text{if } z = z^s_{rare} \neq z^s_{common}, \\ 1, & \text{if } z = z^s_{common} = z^s_{rare}, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Given the subset of outcomes observed from sampling option k, $X_k = \{z_1, z_2, \ldots\}$, the posterior probability of each option state is determined using Bayes rule,

$$p(s|X_k) = \frac{p(X_k|s)p(s)}{\sum_{s' \in S} p(X_k|s')p(s')}, \tag{2}$$

where $p(X_k|s) = \prod_{z \in X_k} p(z|s)$ and $p(s)$ is the prior probability. We assume a flat initial prior over the hypothesis space for each option ($p(s) = 1/|S|$) and that options are independent. Note that this assumption of independence is not applicable to Experiment 3, in which the distribution of options' states in a given problem depended on the condition (e.g., 1 gain/3 losses). The corresponding Bayesian model for Experiment 3 is described in Appendix B.

Given the posterior distribution, the expected reward of option k is found by integrating across possible states,

$$R(k, X) = \sum_{s \in S} p(s|X_k) \cdot \mu(s). \tag{3}$$

Note that when options are independent, the predicted value of an option that has not yet been sampled is $R(k,\{\}) = 0$. Equation 3 thus defines the expected reward from claiming option k given the outcomes observed so far. Finally, the probability of observing a new outcome $z \in \{-200..200\}$ if option k is sampled is given by the marginal probability,

$$p(z|k,X) = \sum_{s \in S} p(z|s) \cdot p(s|X_k). \tag{4}$$

Decision Models

Model 1: Constant stopping probability. The Constant model assumes that people stop exploring according to a constant probability, q. If the player stops, the predicted value of choosing option k is equal to the predicted reward (Equation 3) given the outcomes observed so far (X),

$$V_{choose}(k,X) = R(k,X). \tag{5}$$

Final choices are modeled with a softmax function, such that the probability of stopping and choosing option k is

$$p(\text{choose } k) = q \cdot \frac{exp(V_{choose}(k,X) \cdot \phi)}{\sum_{j \in O} exp(V_{choose}(j,X) \cdot \phi)}. \tag{6}$$

The parameter ϕ controls the individual's sensitivity to predicted value. When $\phi = 0$, options are chosen randomly. As ϕ increases, decisions become increasingly deterministic with respect to the predicted value.

If the player decides *not* to stop they must select an option to explore. The Bayesian model is used to evaluate the benefit of exploring each option in terms of the predicted value of choosing after an additional outcome is observed (referred to as a *preposterior analysis*, see Berger, 1985). The value of sampling option k is defined as the expected maximum option value after having observed an additional outcome z, weighted by the probability of z occurring:

$$V_{sample}(k, X) = \sum_{z \in Z} p(z|k, X) \cdot [\max_{j \in O} V_{choose}(j, X \cup z)]. \tag{7}$$

The probability of sampling option k is again modeled using a softmax function with the same sensitivity parameter ϕ, now multiplied by the probability that the player chose to not stop:

$$p(\text{sample } k) = (1 - q) \cdot \frac{exp(V_{sample}(k, X) \cdot \phi)}{\sum_{j \in O} exp(V_{sample}(j, X) \cdot \phi)}. \tag{8}$$

Thus, under the Constant model stopping decisions are based solely on the value of q and are therefore independent of experienced outcomes. The player then proceeds to select an option based on its predicted value, both for final choices and exploration. The parameters q and ϕ are assumed to be fixed during a game. Note, however, that a player relying on this strategy might respond to competitive pressure by adjusting the probability of stopping across games (e.g., increasing q after losing out to an opponent).

Model 2: Choice-first. Like the Constant model, the Choice-first model assumes that the player first decides whether to stop and choose an option. However, the probability of stopping depends on the current predicted value of the options. If an option is attractive based on previously observed outcomes, the player is more likely to stop and claim it on the current round rather than continue exploring.

The value of choosing and sampling decisions are given by Equations 5 and 7. The probability of choosing option k is

$$p(\text{choose } k) = \frac{exp(V_{choose}(k,X)\cdot\phi + c)}{1+\sum_{j\in O} exp(V_{choose}(j,X)\cdot\phi + c)}, \tag{9}$$

where ϕ is the sensitivity parameter and c is a bias parameter. High values of c correspond to a bias toward stopping even when $V_{choose}(k,X)$ is low. This choice rule is equivalent to a multinomial logit model with the "continue sampling" action as the baseline category (Agresti, 1996).

If the player decides to continue exploring, the choice of which option to sample next is modeled in the same way as in the Constant model. Given that the probability that a player continues sampling is $[1-\sum_{j\in O} p(\text{choose } j)]$, the probability of sampling option k is

$$p(\text{sample } k) = [1-\sum_{j\in O} p(\text{choose } j)]\cdot\frac{exp(V_{sample}(k,X)\cdot\phi)}{\sum_{j\in O} exp(V_{sample}(j,X)\cdot\phi)}. \tag{10}$$

The Choice-first model therefore assumes that players are likely to stop and choose options that have a high predicted value; in the absence of such options they continue to explore. As in the Constant model, sampling decisions are based on the predicted value of observing another outcome from a given option. Across multiple games, a player using this strategy might respond to competitive pressure by increasing their overall bias toward choosing immediately, controlled by the c parameter. Increasing c implies a higher likelihood of stopping to claim options with lower predicted values (e.g., choosing an option that has not yet been sampled even though its predicted value is 0).

Model 3: Tradeoff. Under the Tradeoff strategy the decision maker simultaneously compares the value of an immediate choice versus continuing to explore given the possible actions of an opponent. The player's beliefs about the opponent are represented by two parameters: q_{opp}, the probability that the opponent will stop on each round, and ϕ_{opp}, how deterministic his or her choices are with respect to option value. For instance, high q_{opp} means the opponent is very likely to stop, but low ϕ_{opp} means the person is not likely to choose the most attractive option (in other words, such an opponent is expected to choose "fast and loose"). In contrast, low q_{opp} and high ϕ_{opp} means the opponent is less likely to stop on any given round, but will tend to choose the best option when they stop. These two parameters determine how competition affects the predicted values of both an immediate choice and further exploration.

If an opponent decides to stop, the probability that they choose option k is

$$p(\text{opp chooses } k) = \frac{exp(V(k,X) \cdot \phi_{opp})}{\sum_{j \in O} exp(V(j,X) \cdot \phi_{opp})}. \tag{11}$$

Note that the ϕ_{opp} parameter represents the player's *belief* about how deterministically the opponent will choose with respect to the predicted option values. A high value of ϕ_{opp} reflects a form of pessimism such that the opponent is expected to choose the option with the highest value. Consequently, the value of being the second chooser depends on the likely first choice of the opponent (under the assumption that the best remaining option will be chosen):

$$V_{second}(X) = \sum_{k \in O} p(\text{opp chooses k}) \cdot [\max_{j \in O, j \neq k} V(j,X)] \tag{12}$$

The value of choosing on the current round depends on the opponent's decision as follows. If the opponent decides to continue sampling the player is able to choose first. If a tie occurs because the opponent also decides to stop, the player is assigned the first or second choice with equal probability. Thus, the predicted value of stopping and choosing option k is

$$V_{choose}(k,X) = (1-q_{opp}) \cdot R(k,X) + q_{opp} \cdot [.5 \cdot R(k,X) \; + \; .5 \cdot V_{second}(X)] \tag{13}$$

As in the preceding models, the benefit of exploration is based on the expected outcome of drawing an additional sample and then making a choice. Here, however, that benefit must also be offset by the costs of competition, both on the current round and the next round. For an option k, the value of choosing after observing an additional outcome z is the expected maximum value across available options. These values are integrated according to the probability of each outcome occurring, and multiplied by the probability of the opponent not stopping. If the opponent stops on the current trial (with probability q_{opp}), the player is prevented from exploring further and must choose second. This gives the following value function for sampling option k:

$$V_{sample}(k,X) = (1-q_{opp}) \sum_{z \in Z} p(z|X_k) \cdot [\max_{j \in O} V_{choose}(j, X \cup z)] + q_{opp} \cdot V_{second}(X). \tag{14}$$

Finally, the player simultaneously considers both immediate choices and continued exploration, with the probability of each action defined with the softmax choice function:

$$p(\text{choose } k) = \frac{exp(V_{choose}(k,X) \cdot \phi)}{\sum_{j \in O} exp(V_{choose}(j,X) \cdot \phi) + \sum_{j \in O} exp(V_{sample}(j,X) \cdot \phi)} \tag{15}$$

$$p(\text{sample } k) = \frac{exp(V_{sample}(k,X) \cdot \phi)}{\sum_{j \in O} exp(V_{choose}(j,X) \cdot \phi) + \sum_{j \in O} exp(V_{sample}(j,X) \cdot \phi)}. \tag{16}$$

The Tradeoff model has three parameters. The sensitivity ϕ reflects how deterministically a player acts with respect to predicted value. The remaining parameters represent the player's beliefs about the opponent's behavior: the probability that they stop on each round (q_{opp}) and their sensitivity (ϕ_{opp}). As in the previous models, these parameters are assumed to be fixed within a game. Across multiple games, the effects of competitive pressure on exploration are expected to be mediated by changes in these beliefs about opponents (e.g., increasing q_{opp} after experiencing an opponent stop first).

Model Comparison

Each model was fit to data from Studies 1–3. In addition, a fourth Baseline model was fit that assumed a constant stopping probability (a free parameter q), but otherwise random decisions on every round. For Studies 1 and 2, we divided the data into an early phase (games 1–4) and a late phase (games 5–8) and estimated parameters separately for each phase. Given the findings that sample size decreased across rounds, we tested if this shift was reflected in the difference in estimated parameters between early and late games. For Study 3, parameters were estimated separately for within-pair conditions (1-gain and 3-gains). For the final round of each game we only included the decision of the first chooser (since the intended action of the second chooser was not available).

Models were fit through Bayesian estimation using the *PyMC* Python package (Patil, Huard, & Fonnesbeck, 2010)). For each estimated model, chains were run for 20000 samples with 2000 burn-in samples. Deviance information criterion (DIC) was used to compare models. The prior for stopping probability parameters (q and q_{opp}) was a flat $q \sim Beta(1,1)$ prior. The prior for choice sensitivity parameters (ϕ and ϕ_{opp}) was weakly-informative, $\phi \sim Gamma(1,10)$. The prior for the bias parameter was a Normal distribution centered on zero with high variance, $c \sim Normal(0,50)$. Robustness checks with alternative priors led to convergent results with the settings above.

Results

The resulting DIC values are shown in Table 5. In all studies, the Tradeoff strategy was the best overall model as indicated by the DIC values. Although in some cases the Constant model achieved comparable fits (i.e., both phases of Study 1 and the late phase of Study 2), overall the model comparison offers strong support for the Tradeoff strategy.

The mean posterior estimates and HDIs for the parameters of the Tradeoff strategy are shown in Figure 5. Given the decrease in overall sample sizes observed in Studies 1 and 2, we examined how the estimated parameters changed from early to late games using the posterior distribution of the difference (late − early) for each parameter. Credible differences were assessed based on whether the 95% highest-density interval (HDI) for this distribution excluded zero. There was a credible increase in q_{opp} (the opponent's stopping probability) in all three datasets (Study 1: $M = .1$, $HDI = [.02, .19]$; Study 2, 2 options: $M = .22$, $HDI = [.10, .36]$; Study 2, 4 options: $M = .05$, $HDI = [.01, .08]$). There were no credible differences in the ϕ_{opp} parameter (Study 1: $M = .01$, $HDI = [-.29, .30]$; Study 2, 2 options: $M = .10$, $HDI = [-.21, .44]$; Study 2, 4 options: $M = .11$, $HDI = [-.27, .49]$) or the ϕ parameter (Study 1: $M = .01$, $HDI = [-.08, .10]$; Study 2, 2 options: $M = -.01$, $HDI = [-.05, .04]$; Study 2, 4 options: $M = .01$, $HDI = [-.04, .04]$). These results suggest that the changes in sample size observed across games were driven by an increase in perceived competitive pressure, here represented by the belief about the likelihood of an opponent stopping on any given trial.

We conducted the same comparisons for Study 3, in which the 1-gain condition was associated with increased sample sizes relative to the 3-gain condition. There was a credible increase in q_{opp} in 3-gain games relative to 1-gain games ($M = .15$, $HDI = [.10, .20]$). In addition, there was a credible increase in ϕ in 3-gain games as compared to 1-gain games ($M = .11$, $HDI = [.05, .17]$). There was not a credible difference in the ϕ_{opp} parameter ($M = .12$, $HDI = [-.23, .52]$). The effect of option distribution on sample size observed in Study 3 thus appears to be due, at least in part, to a difference in the perceived competitive pressure in terms of the probability that the opponent will stop (q_{opp}).

One way to compare participants' behavior with the estimated models is to examine the first decision in each round, at which point participants have observed a single outcome. (Comparisons on subsequent rounds are more difficult to visualize given the wide variety of outcome sequences experienced by different participants, even by the second round.) We examined the proportion of participants who made each of four types of decisions on the first round: 1) stop and choose the same option, 2) stop and choose a different option which has not yet been sampled, 3) sample again from the same option, and 4) sample from a different option. The black lines in Figure 6 indicate for each dataset the proportion of each decision as a function of the first observed outcome (binned in order to increase the amount of data in the upper and lower extremes). The dashed line and shaded region indicate the mean and 95% HDI of the posterior predictive distribution of the probability of each action from the estimated Tradeoff model. In general, the model successfully captures the relationship between observed outcomes and participants' decisions on the first round. One notable mismatch is the proportion of 2-option participants who chose to switch to sampling a different option. Whereas the model predicts similar proportions of sampling either option, participants were somewhat more likely to switch to exploring the other option. This may indicate an exploratory "bonus" assigned to options that have not yet been explored (Sutton, 1990).

Discussion

Participants from all three studies were best-described by the compensatory Tradeoff strategy. It weighs the value of further exploration (i.e., drawing another outcome) against that of an immediate, exploitative choice, taking into account the potential costs of a competitor choosing first. When separately fitting behavior in the first and second halves of Studies 1 and 2, we found an increase in the estimated parameter for the probability of an opponent stopping (q_{opp}). This shift reflects the changing value of exploration as competitive pressure increases. As participants experience competition (e.g., losing out to an opponent choosing first), they may adjust this probability upward, causing the predicted reward from continued exploration to decrease relative to that of an immediate choice.

The Tradeoff model provides a parsimonious account of how beliefs about the choice environment and competitors' behavior jointly affect exploration. The model replicates several aspects of the simulation results of Phillips et al. (2014) despite the differences in procedure and choice environments. For instance, given a fixed number of competitors, sample sizes are predicted to increase with the number of options since there is more to be gained from exploration when more options are available. At the same time, fast-acting competitors pose a greater cost that may outweigh the benefits of continued exploration, particularly when the number of options is small.

In addition, the model explains why the effect of competition depends on experienced outcomes and other properties of the choice environment. Consider Phillips et al.'s finding that many choices in their study depended on the valence of the first draw, consistent with a "take-good-enough, otherwise-shift" (TGE) heuristic (Phillips et al., 2014). When the first outcome was positive, participants frequently stopped right away and claimed the same option. When it was negative, they typically switched to either choose or explore the other (unsampled) option. Qualitatively, this heuristic captures some aspects of behavior in the present studies as well. Participants in all studies were likely to stop and choose an option when the first draw was positive (top row of Figure 6). Note, however, that the reaction to negative outcomes differed across studies. In 2-option games, participants frequently responded by immediately choosing the other (unsampled) option, whereas in 4-option games they rarely did so, opting instead to continue exploring a different option. Moreover, in Study 3, participants more frequently stopped to claim an unsampled option after experiencing an extreme negative outcome in the 3-gain condition than in the 1-gain condition. It is unclear how to reconcile these differences with a single heuristic that ignores these variations in the choice environment. Finally, it is worth noting that the TGE heuristic is closely related to the Choice-first model (in that attractive outcomes tend to cause immediate choices) which provided a poorer account of the data.

The success of the Tradeoff model may be related to our participants' knowledge of the choice environment. This environment was also designed to have a simpler probabilistic structure than previous incarnations of the sampling paradigm in order to make the value of exploration more transparent. If people are ignorant of the number or distribution of outcomes ahead of time (as was the case in Phillips et al., 2014), they may be less able to predict the value of continued exploration. This might increase reliance on a strategy in which priority is given to stopping decisions given based on outcomes that have been experienced so far (as in the Choice-first model).

As an exploratory analysis, our approach involved a number of simplifications that could be addressed in further work. Given the small number of games played by each pair of participants, we modeled behavior at the aggregate level, potentially obscuring variability in strategy use across pairs. For instance, some pairs had constant sample sizes that did not vary across games or as a function of observed outcomes. This behavior may be better described by the Constant model if parameters were estimated at the level of individual pairs. In addition, further work is necessary to directly test our proposal for how people learn about competitors across repeated games (i.e., by increasing q_{opp} in response to an opponent stopping first). This would likely benefit from a larger number of games per pair and a larger option set size to permit for a wider range in sample sizes. Finally, the Tradeoff strategy relies on a relatively simple representation of beliefs about competitors (q_{opp} and ϕ_{opp}). People may engage in more sophisticated forms of reasoning in order to evaluate the risks posed by competitors, including expectations about how competitors search (Wilke et al., 2015) or higher levels of iterated reasoning (Ho, Camerer, & Weigelt, 1998; Stahl & Wilson, 1995). The Tradeoff model is a first-order iterated reasoning process (Ohtsubo & Rapoport, 2006) because it assumes that opponents adopt the Constant strategy. Future work could extend the model to investigate how reasoning about others' exploration affects perceived competitive pressure.

General Discussion

Exploration is essential for taming uncertainty across many kinds of decision making environments (Todd, Hills, & Robbins, 2012). Yet reducing uncertainty through exploration rarely comes without costs. Competition for limited resources is one common factor that poses costs for the individual who searches or deliberates too long. Given the ubiquity of competitive pressure, it is important to understand how people perceive and respond to it when making decisions under uncertainty.

In agreement with the results of Phillips et al. (2014), people sharply curtailed their exploration in the face of competition as compared to solitary players (Study 1). Yet exploration also proved sensitive to changes in the environment that affected the degree of competition. Specifically, people collected larger samples when the number of available options increased (Study 2) and thus the ratio of options to competitors became less fierce. One potential explanation for this increased exploration centers on people's beliefs about how long their opponents would explore a larger option set. We examined this further in Study 3 by manipulating knowledge of the option EV distribution (varying the relative proportion of gains and losses), and by extension, the belief about the competitors' propensity to search. Sample sizes increased when losses became more plentiful. Our model-based analysis suggests that this shift resulted from differences in perceived competitive pressure rather than a change in the outcomes experienced (i.e., frequent negative outcomes in 1-gain games).

In contrast to the effects of competitive pressure and option EV distribution, we did not find any impact of on-line payoff feedback on exploration in Study 2. In Study 1, participants received social feedback indicating whether they succeeded in choosing first but did not learn about their decisions' payoffs, potentially causing them to prioritize stopping first. We expected that observing the actual outcomes of their choices might allow people to learn that they were more successful at choosing the *H* option when they experienced a rare outcome. Such insight could potentially counteract the downward pressure on sample size from competitive pressure. However, providing feedback about the consequences of one's own choice had no effect on sample size; furthermore, providing feedback about both participants' choices actually led to slightly lower sample sizes. This raises the possibility that the provision of full feedback, by enabling individuals to compare their performance with their opponents', may amplify the perceived competitive pressure rather than encourage further exploration (see also Hafenbrädl & Woike, 2018).

Learning to Act Fast: A Race to the Bottom

The decline in sample size over the course of repeated games suggests that participants adjusted how much they explored as a result of experiencing competition. In Studies 1 and 2, decreases in sample size were more likely when first choosers had lost out to their opponents in the previous trial. The comparison with the Independent condition in Study 1 demonstrates that this decline did not result from increasing familiarity or practice with the choice environment or mounting opportunity costs imposed by a group experiment. In general, our results suggest a "race to the bottom" that reflects a short-term adaptation to competitive pressure. Under this process, participants may begin the task with high uncertainty about their opponents' behavior and explore options in a manner similar to that of a solitary participant. As participants experience competition, they update their beliefs about their opponents and decrease how much they explore. This repeated interaction leads to a feedback loop within a group of competitors, causing them to converge toward a strategy of minimal exploration.

This type of adaptation has also been found in strategic games in which groups of competitors converge to stable strategies as a result of experience, both over individual and evolutionary time-scales (Avrahami, Güth, Hertwig, Kareev, & Otsubo, 2013; Camerer, 2003; Rapoport, Stein, Parco, & Nicholas, 2003). A recent study by Hintze, Phillips, and Hertwig (2015) illustrates how a minimal exploration strategy emerges when extreme competition is a stable and recurrent property of the ecology. They conducted evolutionary simulations using tasks of a similar nature to the rivals-in-the-dark game, with varying levels of competitive pressure.

Under direct competition, two agents could explore to learn about the value of a common option until deciding between that option and a private alternative of known value (i.e., a sure-thing that was not available to the opponent). In an extreme competition condition similar to that of the current studies, the two agents could sample and claim either of the two options. The strategy that evolved in the direct competition environment was sensitive to environmental variability: agents frequently sampled more than once and the likelihood of continuing to sample increased with outcome variance. In contrast, the strategy that evolved under extreme competition was a minimal one-sample strategy, regardless of the uncertainty in the option value.

Tomlin, Rand, Ludvig, and Cohen (2015) presented a similar set of findings in the context of intertemporal choice. They used a dual-process framework to examine the evolution of strategies that combine a fast, automatic component (i.e., immediate consumption of an entire resource) and a slow, controlled component (i.e., weighing immediate consumption against saving resources for the future). They assumed that when two agents compete for the same resource, an agent following an automatic strategy acts faster than an opponent relying on a controlled strategy. In the absence of competition, a controlled strategy is advantageous because it enables flexible consumption based on an agent's current state of energy and the availability of resources in the environment. In highly competitive environments, however, the stable evolutionary strategy is one with a high propensity for fast, automatic responses that reduce the chance of losing out to others.

These lines of work demonstrate how learning to act fast, even when faced with high degree of uncertainty about the quality of the options, can be adaptive. The key is the recurrent presence of extreme competition. Accordingly, people's willingness to explore in social settings may depend on the kinds of competition they have experienced in the past. Recent work has suggested that manifestations of seemingly impulsive choice may in fact reflect adaptation to stressful or highly uncertain social environments (Frankenhuis, Panchanathan, & Nettle, 2016; Kidd, Palmeri, & Aslin, 2013). In a similar vein, experiencing intense competition for resources in the past (e.g., due to socioeconomic background or experience in a highly competitive industry) may lead to less exploration even in contexts in which competitive pressure is eased.

Implications for Other Social Environments

We have focused on a relatively austere competitive environment. Individuals were forced to stop and choose an option when their opponents terminated search, even when multiple additional options were available. In many of the real-world examples of competitive choice that we have discussed, people are not mandated to stop exploring at the same time as their competitors. However, it is often the case that continuing to search after opponents have stopped incurs additional costs, as when opponents gain a competitive edge from their choice which affects later interactions (e.g., when a competing company hires an star employee that makes it easier to attract additional talent). Although such *first-mover advantages* are typically examined in the context of organizational decision making (e.g., Lieberman & Montgomery, 1988), similar costs may imply high competitive pressure even when individuals are free to explore independently of others.

Our studies also provided participants with scant information about opponents' behavior, whereas real-world competition often features richer social interactions. Research on behavioral ecology has examined competitors' use of *public information*, defined as observations of competitors' choices that are used to assess the quality of a resource (Danchin, Giraldeau, Valone, & Wagner, 2004). For instance, one advantage of foraging in a group (rather than alone) is that the individual can learn about the distribution of resources by observing other group members' search behavior. Although patches of resources are depleted more quickly due to consumption by competitors, groups of foragers can use public information to better discern when a patch is exhausted and it is time to explore further afield (Valone & Templeton, 2002). Since the rivals-in the-dark game separates an initial phase of exploration from a single exploitative choice, public information about the final choice could have no effect on exploration in the competitive conditions. In the Independent condition in Study 1, the second chooser in each game did observe which option was selected by the first chooser, but could also continue exploring to learn about either option. Thus, public information about choices is unlikely to have affected behavior in the present studies. We would expect it to play a greater role in competitive environments that involve ongoing exploitation of a large number of options (Goldstone, Ashpole, & Roberts, 2005).

In addition to seeing others' final choices, public information during exploration may offer additional benefits. Observing which options an opponent samples (and the frequencies of sampling) can provide a signal of their quality even if the actual outcomes are not public. For example, if an opponent samples an option once and immediately switches to explore a different option, one might infer that they did not experience an especially favorable outcome. Finally, observing the outcomes of others' exploration would lead to obvious benefits in terms of estimating option values as well as predicting opponents' decisions to stop and choose. To take an example from the domain of mate search, people exhibit mate-copying behavior such that evaluations of potential mates are influenced by observations of their interactions during speed dating (Place, Todd, Penke, & Asendorpf, 2010). An important consideration for future work is whether competitors prolong exploration when it offers these additional opportunities for social learning.

Conclusions

For human decision makers, competitive pressure is both ubiquitous and heterogeneous. Different environments bring unique tradeoffs between the costs of losing out to opponents and the benefits of exploring to reduce uncertainty. Our results suggest that people adapt their exploration based on the features of a novel and unfamiliar competitive environment and as a result of experiencing competition for finite resources (consistent with a “race to the bottom” over the course of repeated play). These findings highlight the need to consider how the social dimension of experience, including both past and present exposure to competitive pressure, affects how people explore when making decisions under uncertainty.

Appendix A: Simulating $p(H)$ based on ideal observer model

An ideal observer model was developed using the Bayesian updating process described in the Belief Updating section above, with the additional assumption that the observer deterministically chooses the option k with the higher predicted value. Ten-thousand sets of eight two-option problems were randomly generated, subject to the constraints that the EV of the two options differed by more than 25, the summed EV of the best options was less than 200, and the summed EVs of the worst options was greater than -100 (consistent with the procedure of Study 1). We first evaluated how the proportion of H choices in 2-option problems depended on the subset of outcomes experienced by the learner. For each problem we found the predicted choice after observing different subsets of outcomes corresponding to each cell of Table 1, with the assumption that each outcome is experienced only once. We then calculated the proportion of problems for which the model chose option H. Note that there is a small proportion of problems where a rare outcome falls within the same range as the common outcomes, and the observer will be uncertain about its probability based on a single observation. Thus, the results shown in Table 1 do not assume that the observer knows whether a particular outcome is rare or common, but simply reflects $p(H)$ given at most a single observation of each outcome type for a given problem.

Our next goal was to assess how $p(H)$ changes with increasing sample size. For each problem we randomly generated 100 sets of observations of sample size N, assuming an equal likelihood of sampling from options H and L. Model performance for each value of N was measured as the frequency of H choices, averaged across runs and problems. The resulting $p(H)$ is shown by the dashed line in Figure 1C.

Appendix B: Bayesian model for dependent options (Study 3)

In Study 3, option states were generated according to the condition (1 gain/3 losses; 3 gains/1 loss). As a result, each observed outcome conveys information about the state of the sampled option as well as the remaining options. For example, in the 1-gain condition, observing a positive outcome from one option leads to a decreased predicted value of the remaining three options (since they are likely to be losses). In the following we describe the Bayesian model that accounts for this dependency between options.

The joint hypothesis space S is comprised of all possible combinations of states across four options, $S = \{(z_c^a, z_r^a, z_c^b, z_r^b, z_c^c, z_r^c, z_c^d, z_r^d) : z_c^k \in Z_{common}, z_r^k \in Z_{rare}\}$. Given a state $s \in S$, each option k is associated with a reward equal to the expected value, $R_k(s) = 0.8 \cdot z_c^k + 0.2 \cdot z_r^k$. In Study 3, the condition specifying the proportion of gains and losses determines the prior distribution. Let S_{1gain} be the subset of states for which three options have negative expected values and one option has positive expected value, while S_{3gain} is the subset with the reversed proportion. In the 1-gain condition, the prior probability is then uniformly distributed over states with a single gain,

$$p(s) = \begin{cases} \frac{1}{|S_{1gain}|}, & \text{if } s \in S_{1gain}, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

The prior probability for the 3-gain condition is defined analogously by replacing S_{1gain} with S_{3gain} in Equation 17. The likelihood function is now defined with respect to the individual option:

$$p_k(z|s) = \begin{cases} .8, & \text{if } z = z_c^k \neq z_r^k, \\ .2, & \text{if } z = z_r^k \neq z_c^k, \\ 1, & \text{if } z = z_c^k = z_r^k, \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

Given the subset of outcomes observed so far from sampling each option k, $X_k = \{z_1, z_2, \ldots\}$, the posterior probability of each state is determined using Bayes rule,

$$p(s|X) = \frac{p(X|s)p(s)}{\sum_{s' \in S} p(X|s')p(s')}, \tag{19}$$

where $p(X|s) = \prod_{k \in O} \prod_{z \in X_k} p_k(z|s)$ and $p(s)$ is the prior as determined by the experimental condition.

Given the posterior distribution over option states, the expected reward of option k is found by integrating across possible states,

$$V(k,X) = \sum_{s \in S} p(s|X) \cdot R_k(s). \tag{20}$$

and the probability of observing a new outcome $z \in \{-200..200\}$ if option k is sampled is given by the marginal probability,

$$p(z|k,X) = \sum_{s \in S} p_k(z|s) \cdot p(s|X). \tag{21}$$

References

Agresti, A. (1996). *Categorical data analysis* (Vol. 990). New York: John Wiley & Sons.

Avrahami, J., Güth, W., Hertwig, R., Kareev, Y., & Otsubo, H. (2013). Learning (not) to yield: An experimental study of evolving ultimatum game behavior. *The Journal of Socio-Economics*, *47*, 47–54.

Berger, J. O. (1985). Preposterior and sequential analysis. In *Statistical decision theory and bayesian analysis* (pp. 432–520). Springer.

Bröder, A. (2000). Assessing the empirical validity of the" take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1332.

Busemeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology*, *32*(2), 91–134.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Cohen, J., McClure, S., & Yu, A. (2007). 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*, *362*(1481), 933.

Danchin, É., Giraldeau, L.-A., Valone, T. J., & Wagner, R. H. (2004). Public information: from nosy neighbors to cultural evolution. *Science*, *305*(5683), 487–491.

Dickson, P. R. (1992). Toward a general theory of competitive rationality. *The Journal of Marketing*, 69–83.

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments.

Academy of Management journal, *32*(3), 543–576.

Frankenhuis, W. E., Panchanathan, K., & Nettle, D. (2016). Cognition in harsh and unpredictable environments. *Current Opinion in Psychology*, *7*, 76–80.

Frey, R., Mata, R., & Hertwig, R. (2015). The role of cognitive abilities in decisions from experience: Age differences emerge as a function of choice set size. *Cognition*, *142*, 60–80.

Fu, W., & Gray, W. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology*, *52*(3), 195–242.

Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press, Inc.

Goldstone, R. L., Ashpole, B. C., & Roberts, M. E. (2005). Knowledge of resources and competitors in human foraging. *Psychonomic Bulletin & Review*, *12*(1), 81–87.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, *118*(4), 523.

Gray, W., & Fu, W. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, *28*(3), 359–382.

Gureckis, T. M., Martin, J., McDonnell, J. V., Rich, A. S., Markant, D., Coenen, A., … Chan, P. (2015). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, 1–14. doi: 10.3758/s13428-015-0642-8

Hafenbrädl, S., & Woike, J. K. (2018). Competitive escalation and interventions. *Journal of Behavioral Decision Making*, *31*(5), 695–714.

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description–experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, *21*(5), 493–518.

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

Hertwig, R., & Hoffrage, U. (2013). *Simple heuristics in a social world*. Oxford University Press.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*(2), 225–237.

Hills, T. T., Noguchi, T., & Gibbert, M. (2013). Information overload or search-amplified risk? set size and order effects on decisions from experience. *Psychonomic bulletin & review*, *20*(5), 1023–1031.

Hintze, A., Phillips, N., & Hertwig, R. (2015). The janus face of darwinian competition. *Scientific reports*, *5*.

Ho, T.-H., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental" p-beauty contests". *The American Economic Review*, *88*(4), 947–969.

Kidd, C., Palmeri, H., & Aslin, R. N. (2013). Rational snacking: Young children's decision-making on the marshmallow task is moderated by beliefs about environmental reliability. *Cognition*, *126*(1), 109–114.

Lejarraga, T., & Hertwig, R. (2016). How the threat of losses makes people explore more than the promise of gains. *Psychonomic Bulletin & Review*, 1–13.

Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition*, *124*(3), 334–342.

Lieberman, M. B., & Montgomery, D. B. (1988). First-mover advantages. *Strategic Management*

Journal, *9*(S1), 41–58.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical surveys*. Wiley New York.

Markant, D., Settles, B., & Gureckis, T. M. (2016). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*, *40*(1), 100–120. doi: 10.1111/cogs.12220

Mehlhorn, K., Ben-Asher, N., Dutt, V., & Gonzalez, C. (2014). Observed variability and values matter: Toward a better understanding of information search and decisions from experience. *Journal of Behavioral Decision Making*, *27*(4), 328–339.

Nelson, J. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *114*(3), 677.

Ohtsubo, Y., & Rapoport, A. (2006). Depth of reasoning in strategic form games. *The Journal of Socio-Economics*, *35*(1), 31–47.

Patil, A., Huard, D., & Fonnesbeck, C. J. (2010). Pymc: Bayesian stochastic modelling in python. *Journal of statistical software*, *35*(4), 1.

Payne, J. W., Bettman, J. R., & Luce, M. F. (1996). When time is money: Decision behavior under opportunity-cost time pressure. *Organizational Behavior and Human Decision Processes*, *66*(2), 131–152.

Phillips, N. D., Hertwig, R., Kareev, Y., & Avrahami, J. (2014). Rivals in the dark: How competition influences search in decisions under uncertainty. *Cognition*, *133*(1), 104–119.

Place, S. S., Todd, P. M., Penke, L., & Asendorpf, J. B. (2010). Humans show mate copying after observing real mate choices. *Evolution and Human Behavior*, *31*(5), 320–325.

Rapoport, A., Stein, W. E., Parco, J. E., & Nicholas, T. E. (2003). Equilibrium play and adaptive learning in a three-person centipede game. *Games and Economic Behavior*, *43*(2), 239–265.

Rapoport, A., & Tversky, A. (1970). Choice behavior in an optional stopping task. *Organizational Behavior and Human Performance*, *5*(2), 105–120.

Ratchford, B. T. (1982). Cost-benefit models for explaining consumer choice and information seeking behavior. *Management Science*, *28*(2), 197–212.

Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, *127*(2), 258–276.

Stahl, D. O., & Wilson, P. W. (1995). On playersâĂš models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*(1), 218–254.

Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy*, 213–225.

Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In *Proceedings of the seventh international conference on machine learning* (pp. 216–224).

Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. OUP USA.

Todd, P. M., Hills, T. T., & Robbins, T. W. (2012). *Cognitive search: Evolution, algorithms, and the brain*. MIT press.

Tomlin, D., Rand, D. G., Ludvig, E. A., & Cohen, J. D. (2015). The evolution and devolution of cognitive control: The costs of deliberation in a competitive world. *Scientific reports*, *5*.

Valone, T. J., & Templeton, J. J. (2002). Public information for the assessment of quality: a widespread social phenomenon. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1427), 1549–1557.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*.

Princeton, NJ: Princeton University Press.

Wilke, A., Minich, S., Panis, M., Langen, T. A., Skufca, J. D., & Todd, P. M. (2015). A game of hide and seek: Expectations of clumpy resources influence hiding and searching patterns. *PloS One*, 1–20. doi: 10.1371/journal.pone.0130976

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*, *143*(6), 2074.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological bulletin*, *144*(2), 140.

Table 1

Probability of obtaining H based on partial outcome experience from two options H and L.

		Observed rare outcomes		
		None	H ∣ L	H & L
	None	.5	.81	.99
Observed common outcomes	H ∣ L	.59	.83	.99
	H & L	.62	.85	1

Table 2

Estimated fixed effects from negative binomial regression model of sample size

	β	95%-l	95%-u	Wald *z*	*p*
Study 1					
Intercept	0.89	0.67	1.10	8.17	< .001
Condition (Competitive)	-0.91	-1.32	-0.51	-4.45	**< .001**
Trial (1-8)	0.05	0.02	0.07	4.09	**< .001**
Condition × Trial	-0.16	-0.22	-0.11	-5.61	**< .001**
Study 2					
Intercept	-0.40	-.67	-0.15	-3.03	0.01
Feedback (Partial)	0.09	-0.30	0.50	0.48	0.63
Feedback (Both)	-0.02	-0.31	0.28	-0.14	0.89
Number of options (4)	1.05	0.69	1.42	5.75	**< .001**
Trial (1-8)	-0.11	-0.13	-0.09	-10.46	**< .001**
Feedback (Partial) × Number of options	-0.16	-0.74	0.42	-1.13	0.59
Feedback (Both) × Number of options	-0.27	-0.74	0.20	-0.53	0.26
Study 3					
Intercept	0.21	-0.09	0.49	1.46	0.15
Number of gains (3)	-0.25	-0.40	-0.10	-3.18	**.001**
Trial (1-8)	-0.02	-0.05	0.01	-1.26	0.21

Table 3

Estimated fixed effects from logistic regression model of decrease in sample size across trials in competitive conditions

	β	95%-l	95%-u	Wald *z*	*p*
Study 1					
Intercept	-1.07	-1.78	-0.42	-3.11	< .001
Trial (2 – 8)	0.00	-0.18	0.19	0.03	.98
Slower_{t-1}	1.21	0.32	2.13	2.63	**.009**
Trial × Slower_{t-1}	-0.30	-0.57	-0.04	-2.24	**.03**
Study 2					
Intercept	-1.08	-1.40	-0.77	-6.78	< .001
Feedback (Partial)	0.01	-0.25	0.27	0.05	.96
Feedback (Both)	-0.01	-0.27	0.24	-0.11	.91
Number of options (4)	0.64	0.42	0.85	5.81	**< .001**
Trial (2 – 8)	-0.11	-0.18	-0.04	-2.94	**.003**
Slower_{t-1}	0.82	0.45	1.19	4.36	**< .001**
Trial × Slower_{t-1}	-0.12	-0.23	-0.01	-2.14	**.03**
Study 3					
Intercept	-0.88	-1.43	-0.34	-3.17	.002
Number of gains (3)	0.40	-0.01	0.81	1.90	.06
Trial (2 – 8)	-0.02	-0.16	0.12	-0.28	.78
Slower_{t-1}	0.45	-0.27	1.18	1.23	.22
Trial × Slower_{t-1}	-0.11	-0.32	0.10	-1.06	.29

Table 4

Estimated fixed effects from logistic regression on probability of choosing H.

	β	95%-l	95%-u	Wald *z*	*p*
Study 1					
Intercept	0.65	0.28	1.04	3.38	< .001
Condition (Competitive)	0.20	-0.39	0.79	0.65	0.51
Trial (1-8)	0.06	-0.03	0.14	1.38	0.18
Condition × Trial	-0.16	-0.29	-0.02	-2.30	**0.02**
Observed 1+ rare outcomes	0.93	0.61	1.26	5.71	**< .001**
Study 2					
Two options					
Intercept	0.34	-0.05	0.87	1.48	0.14
Feedback (Partial)	-0.37	-0.98	0.23	-1.12	0.26
Feedback (Full)	-0.39	-0.98	0.23	-1.25	0.21
Trial (1-8)	0.02	-0.09	0.12	0.34	0.74
Feedback (Partial) × Trial	0.07	-0.11	0.18	0.87	0.38
Feedback (Full) × Trial	0.04	-0.11	0.18	0.50	0.62
Observed 1+ rare outcomes	1.04	0.49	1.22	6.48	**< .001**
Four options					
Intercept	-0.37	-0.81	0.06	-1.62	0.10
Feedback (Partial)	-0.17	-0.63	0.50	-0.54	0.59
Feedback (Full)	-0.07	-0.63	0.50	-0.24	0.81
Trial (1-8)	-0.09	-0.18	0.01	-1.80	0.07
Feedback (Partial) × Trial	0.10	-0.14	0.13	1.34	0.18
Feedback (Full) × Trial	-0.01	-0.14	0.13	-0.08	0.94
Observed 1+ rare outcomes	0.79	0.48	1.11	5.79	**< .001**
Study 3					
Intercept	-0.61	-1.07	-0.17	-2.66	.01
Number of gains (3)	-0.55	-0.97	-0.14	-2.60	**0.01**
Trial (1-8)	0.00	-0.09	0.09	0.05	0.96
Observed 1+ rare outcomes	1.02	0.61	1.45	4.84	**< .001**

Table 5

DIC values from model comparison

	Study 1 (Competitive)		Study 2 (2 options)		Study 2 (4 options)		Study 3	
Model	early	late	early	late	early	late	1-gain	3-gains
Baseline	1549	1234	3438	2563	8931	7598	4088	3445
Constant	1398	1073	3143	2295	8676	7118	3987	3242
Choice-first	1467	1155	3149	2355	7969	6558	3626	3300
Tradeoff	**1394**	**1070**	**3113**	**2291**	**7709**	**6333**	**3508**	**2924**

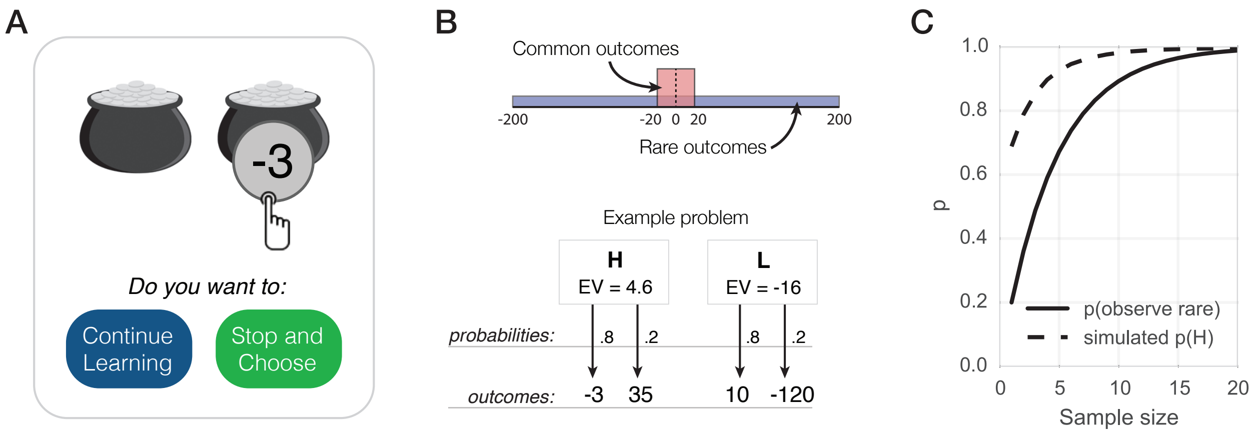

Figure 1. **A:** On each round of the rivals-in-the-dark game the respondent clicks on an option and observes a randomly generated outcome, then decides whether to continue sampling or to stop and choose one of the options. **B:** Binary outcomes for each option were generated by sampling a *common* outcome from a uniform distribution bounded by −20 and 20, and sampling a *rare* outcome from a uniform distribution bounded by −200 and 200. Common and rare outcomes occurred with fixed probabilities of .8 and .2, respectively. An illustrative two-option problem is shown at the bottom, with corresponding outcomes, probabilities, and EVs. **C:** Probability of experiencing at least one rare outcome as a function of sample size (black line). Based on simulated observation sets of varying sample size, the mean performance of a Bayesian ideal observer begins at approximately .7 for a sample size of 1 and approaches perfect accuracy for sample sizes larger than 10 (dashed line).

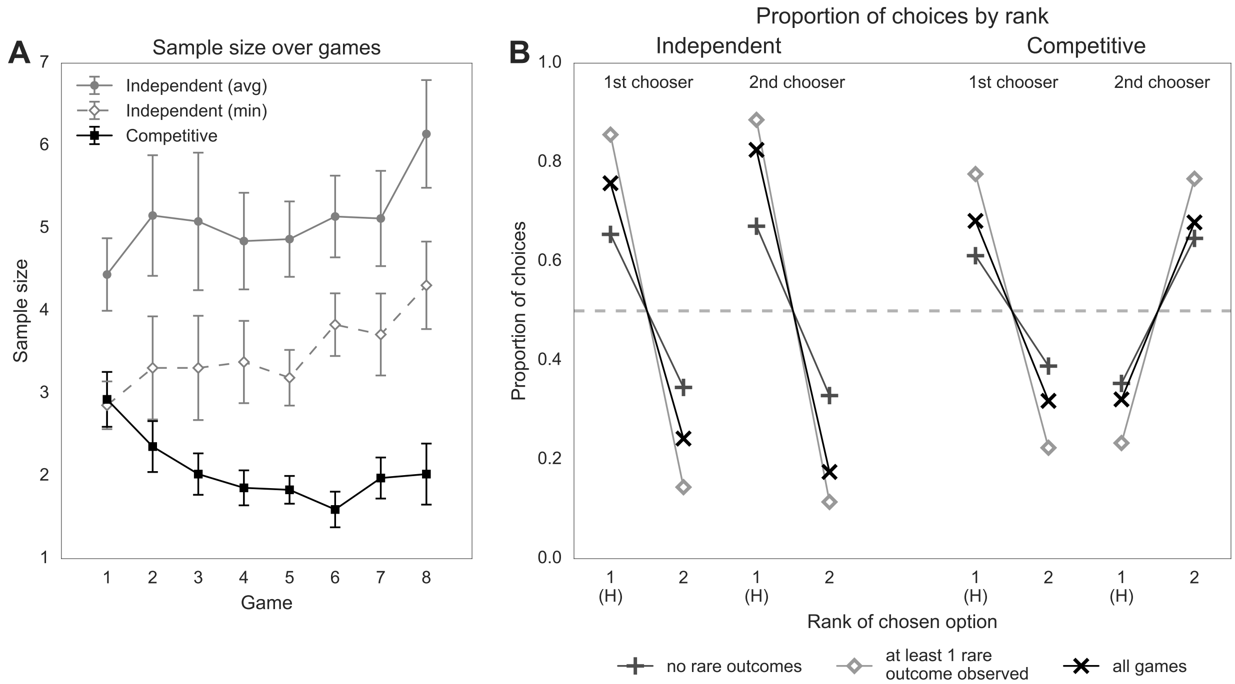

Figure 2. Experiment 1 results. **A:** Mean sample size for the Competitive condition (black line), all participants in the Independent condition (solid gray line), and for the first stoppers in the Independent condition (dashed gray line). Error bars indicate standard errors. **B:** Proportion of games in which each option was chosen, separated by condition and choice order (1st or 2nd chooser).

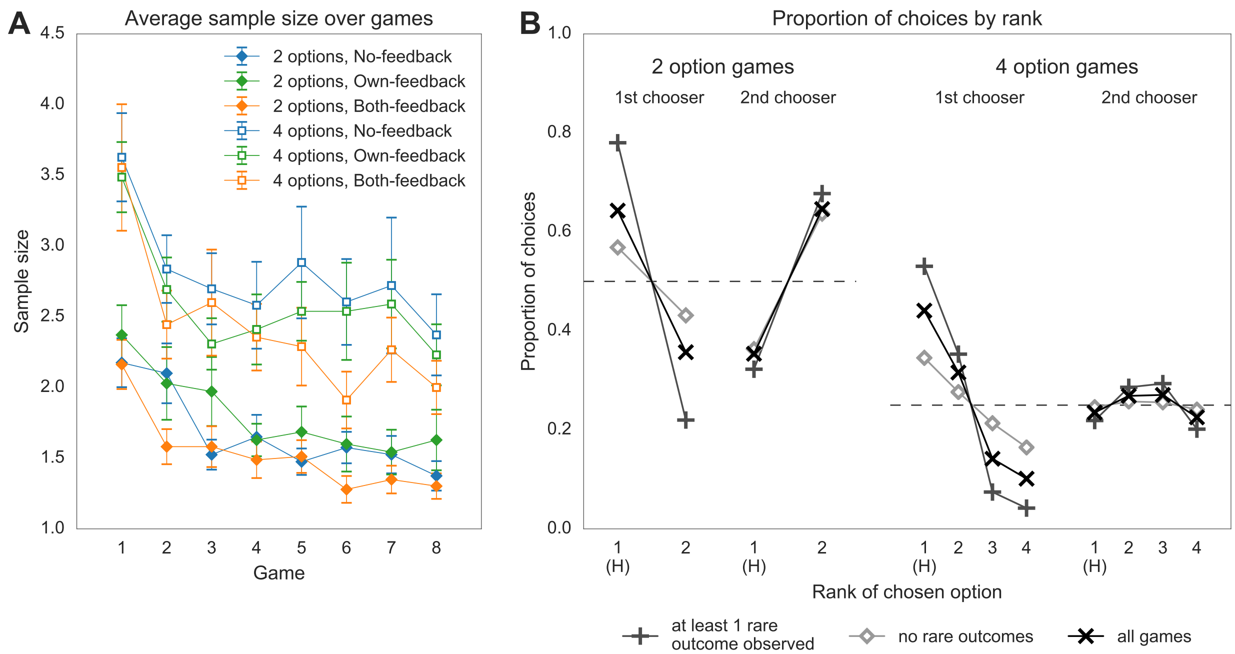

Figure 3. **A:** Mean sample size for each condition in Experiment 2. **B:** Choice proportions by option rank for the first and second choosers in Experiment 2.

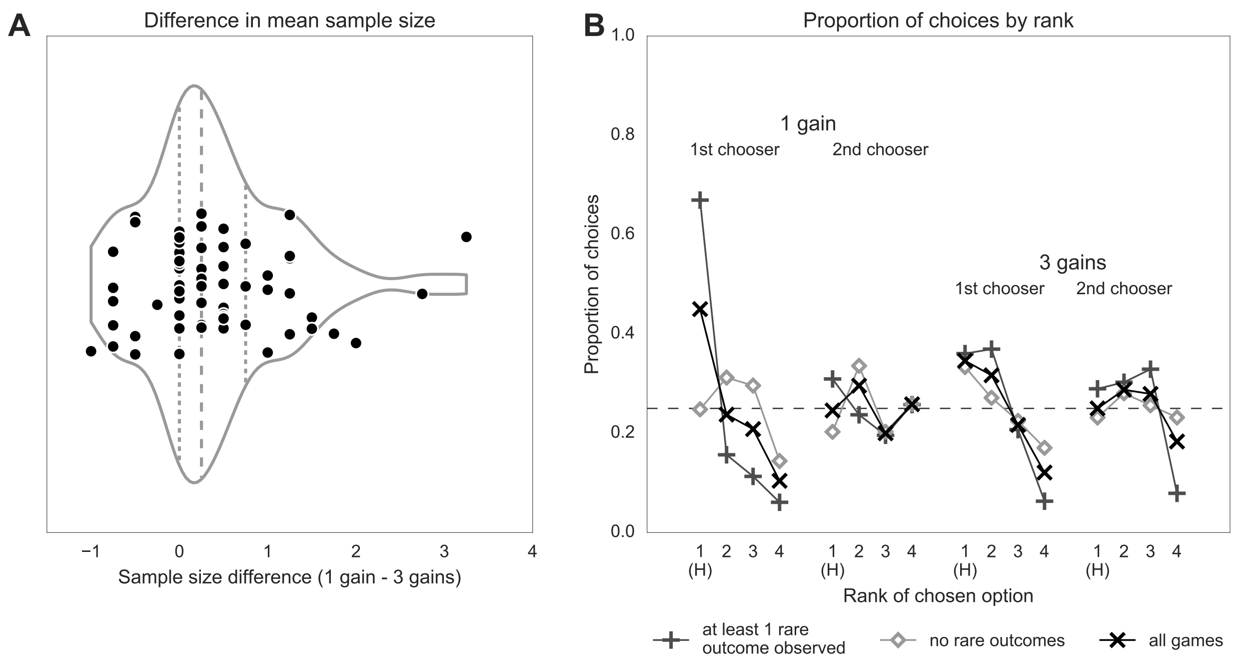

Figure 4. Experiment 3 results. **A:** Within-pair differences in mean sample size between the 1-gain and 3-gains conditions. **B:** Choice performance.

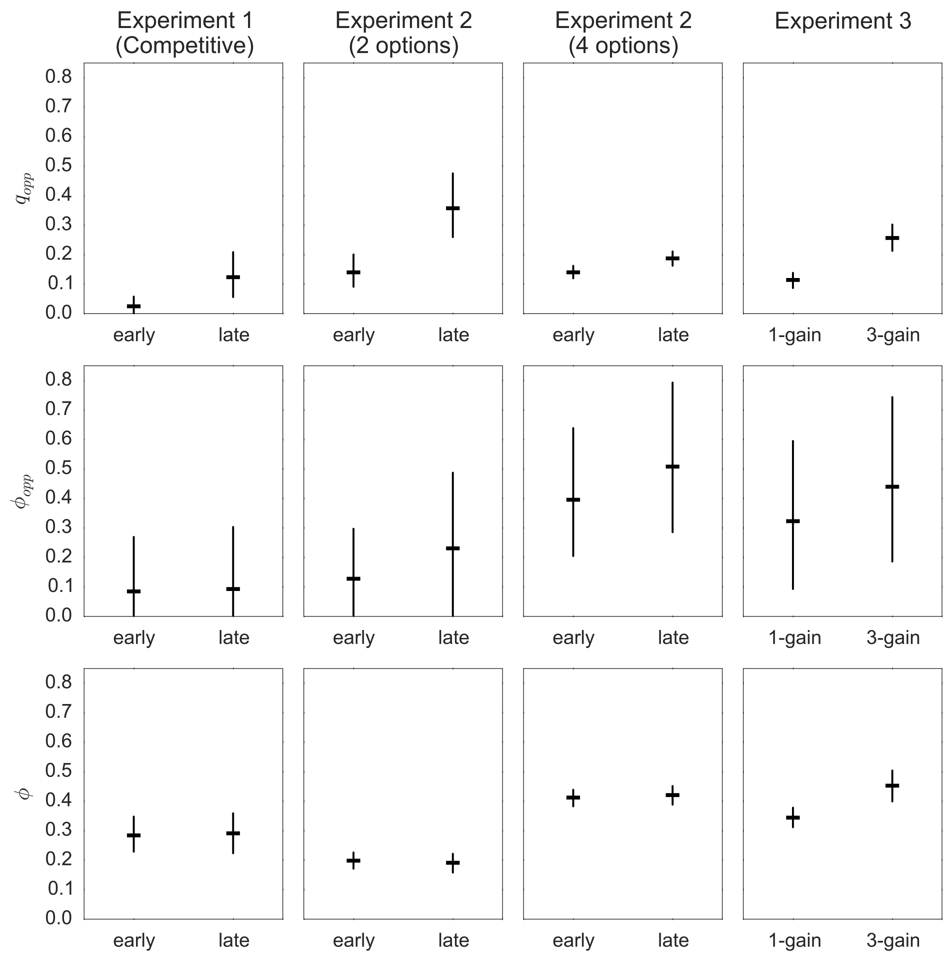

Figure 5. Mean parameter values (horizontal lines) and highest density intervals (vertical lines) from estimated Tradeoff model for each dataset. Parameters were estimated separately for early games (1–4) and late games (5–8) from each experimental condition.

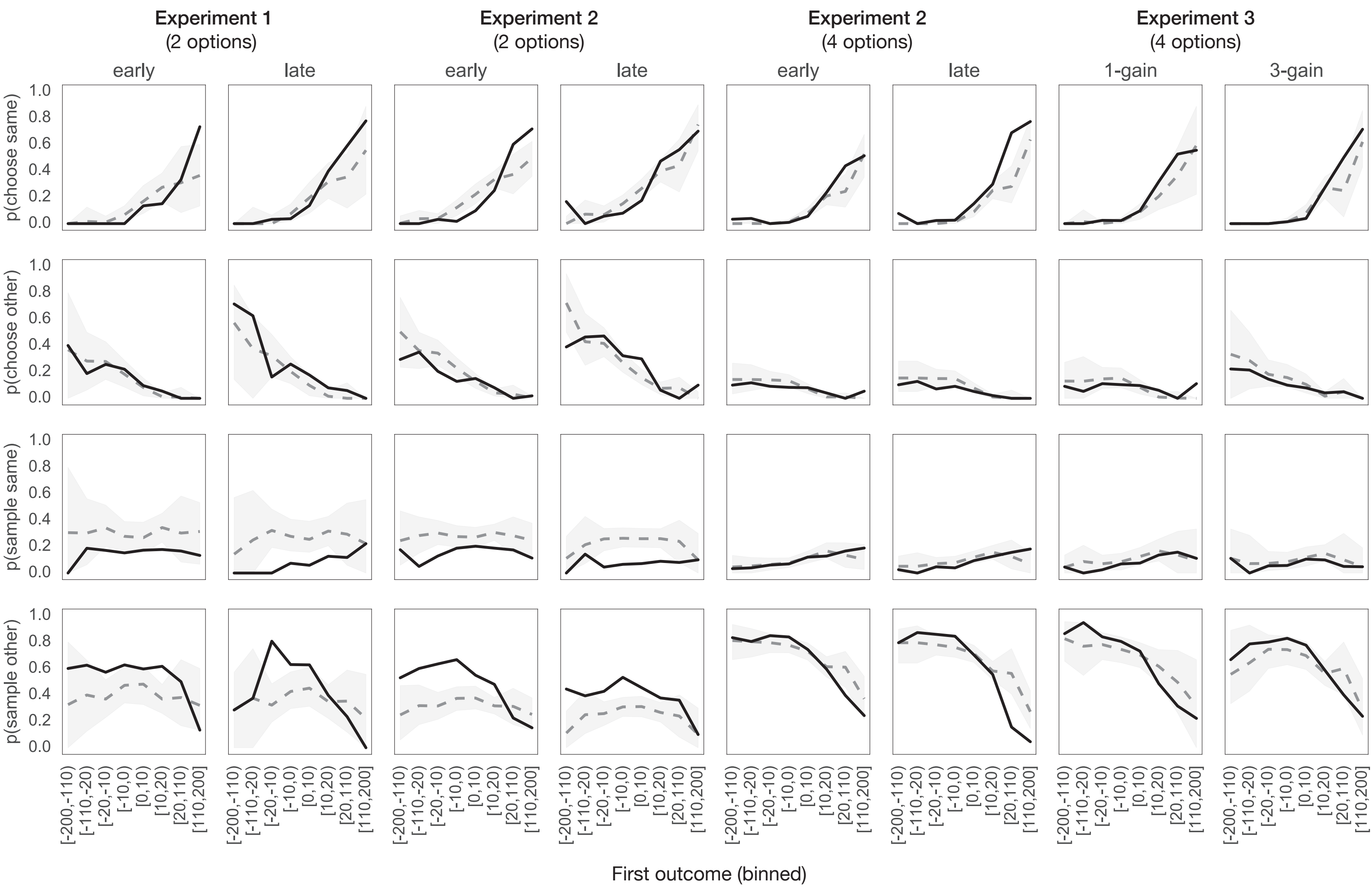

Figure 6. Comparison of participants' actions and predictions of the Tradeoff model on the first round, as a function of the first observed outcome. Black lines indicate the proportion of rounds in which participants chose each action (*sample same*, *sample other*, *choose same*, *choose other*). Gray lines and regions indicate the mean and 95% HDI of the probability of each action based on posterior simulation from the estimated Tradeoff model.