Exploration and exploitation during information search and consequential choice

Cleotilde Gonzalez[1] and Varun Dutt[2]

[1]Dynamic Decision Making Laboratory, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA and [2]School of Computing and Electrical Engineering and School of Humanities and Social Sciences, Indian Institute of Technology Mandi, India

Before making a choice we often search and explore the options available. For example, we try clothes on before selecting the one to buy and we search for career options before deciding a career to pursue. Although the exploration process, where one is free to sample available options is pervasive, we know little about how and why humans explore an environment before making choices. This research contributes to the clarification of some of the phenomena that describe how people perform search during free sampling: we find a gradual decrease of exploration and, in parallel, a tendency to explore and choose options of high value. These patterns provide support to the existence of learning and an exploration-exploitation tradeoffs that may occur during free sampling. Thus, exploration in free sampling is not led by the purely epistemic value of the available options. Rather, exploration during free sampling is a learning process that is influenced by memory effects and by the value of the options available, where participants pursue options of high value more frequently. These parallel processes predict the consequential choice.

Keywords: *choice, decisions from experience, exploration-exploitation, sampling, instance-based learning theory*

An important aspect of decision-making in many daily situations involves a process of exploration of the available options before making a choice for real. Such is the case when we search for information on the web before making a purchase (Pirolli & Card, 1999), when we search around for possible partners before making a dating selection (Todd, Penke, Fasolo, & Lenton, 2007), and when a radiologist examines a scan of a patient for possible diagnosis before deciding the treatment (Wolfe, 2012). Despite the relevance of the exploration process in many naturalistic tasks, we know relatively little about how and why humans explore an environment and how the information obtained from exploration is used in making choices. This research contributes to clarifying some of the aspects of search and exploration in experiential binary choice. In principle, a rational explorer should sample all available options for as long as possible before making a choice given that new information may be collected from exploration, which is expected to lead to better choices. However, previous studies involving free sampling in binary choice reveal at least five patterns of exploration that do not conform to the rational explorer: (1) people rely on surprisingly small samples (Hertwig & Pleskac, 2008, 2010); (2) they tend to sample more when higher and more variable pay-

offs are involved (Hau, Pleskac, Kiefer, & Hertwig, 2008; Mehlhorn, Ben-Asher, Dutt, & Gonzalez, 2014); (3) they follow generally two exploration policies (piecewise or comprehensive strategies) (Hills & Hertwig, 2010); (4) they reduce their rate of exploration over time (Gonzalez & Dutt, 2011, 2012); and (5) they tend to choose the option that they sampled more often (Gonzalez & Dutt, 2012). The main contribution of the current research is the clarification of the relationship between the rate of exploration over time and the tendency to explore and choose the high value option that would lead to the best result. In a binary choice task with free sampling, we demonstrate that a reduction in exploration occurs in parallel with a tendency to select an option with the higher experienced mean more often, regardless of the exploration policy that participants take.

Free sampling and the exploration-exploitation tradeoff

In the study of decisions from experience, researchers have developed an experimental paradigm to study the process of exploration and subsequent choice in a binary task. The paradigm, *sampling paradigm* (see Figure 1), provides a way for participants to explore the two options freely, for as long as they desire and in the order they *desire*, before making one choice for real (Camilleri & Newell, 2011; Hertwig & Erev, 2009; Rakow & Newell, 2010). Although most of the studies in this paradigm have concentrated on highlighting the choice after exploration to contrast with traditional choice from description (e.g., Hertwig, Barron, Weber, & Erev, 2004), this paradigm opens a window for investigating the behavior, processes, and strategies that people pursue during exploration, before making a choice.

The exploration rate in this task has been found to decrease over an increasing number of repeated samples (Gonzalez & Dutt, 2011, 2012; Teodorescu & Erev, 2014); and, people tend to choose the option that they sampled more frequently and more recently (Gonzalez & Dutt, 2011, 2012). A few studies have suggested that the decrease in exploration rate might be related to the process of discovering an option that maximizes outcomes (Gonzalez & Dutt, 2011, 2012); while some find a more extreme effect: that a decrease of exploration occurs even when it is most optimal to keep exploring (Teodorescu & Erev, 2014).

Hills and Hertwig (2012) questioned the robustness of the observation in Gonzalez and Dutt (2011) about the reduction of exploration rate over time and their suggestion

Corresponding author: Cleotilde Gonzalez, Dynamic Decision Making Laboratory, Carnegie Mellon University, Pittsburgh, PA 15213, USA. e-mail: coty@cmu.edu
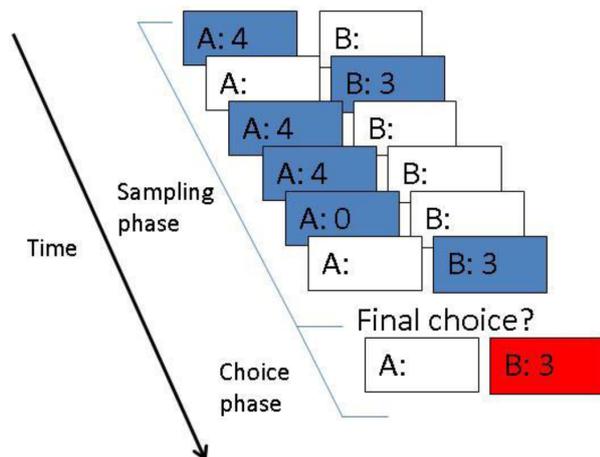
Figure 1. The sampling paradigm of decisions from experience. During the sampling phase people select options freely. By selecting an option an outcome is drawn from a distribution, presented as a result. The figure shows a problem of choice between two options A: a .8 chance of earning $4 and .2 chance of earning $0; and B: Earning $3 for sure. Participants first sample the two options A and B to discover their values and once they are satisfied with the information they choose one of the two options (A or B) for real.

that participants explore options that corresponded to the highest value (Sampling-H). The heart of their argument is that in the sampling paradigm, an impression of reduced exploration over time (alternation rate, A-rate, in binary choice) is produced by an inverse relationship between the sample size and the A-rate, and the aggregation of participants with different sample sizes. Gonzalez and Dutt (2012) showed that the reduction of A-rate during sampling occurs even when sample length is controlled for and that people tend to explore and choose according to the value of the options. They argued that a decrease of exploration in the sampling paradigm might be related to an implicit goal of discovering which of the two options maximizes rewards, suggesting that an exploration-exploitation tradeoff may be occurring during free sampling.

What happens during free sampling is still unclear. One possibility is that exploration is simply random (Hertwig & Erev, 2009; Rakow & Newell, 2010). The assumption of a random search is very reasonable for a rational explorer that wants to maximize the information obtained, and it is very commonly used in a large variety of cognitive models attempting to account for the choice after sampling (see Gonzalez and Dutt, 2011 for a review of these models). However, this assumption of randomness does not explain the patterns of exploration found in human data (Fiedler, 2000; Fiedler & Kareev, 2006; Gonzalez & Dutt, 2011). A random sampling assumption would presuppose a stability of other factors like learning from the sampling process and being influenced by memory effects (e.g., frequency and recency of experienced outcomes). Some argue that in the sampling paradigm there cannot be exploration-exploitation tradeoffs because the sampling process is separated from choice and it can only be used to obtain information without any concerns about costs and rewards but with the only purpose of informing the consequential choice after sampling (Hills & Hertwig, 2012). Psychologists would generally suggest that more informed decisions are a result of larger sample sizes (i.e., the value of information increases with more samples) (Fiedler, 2000). Thus, a strong and robust finding that people draw small

samples (a median number between 11 and 19 times) before making a choice (e.g., Gonzalez & Dutt, 2011; Hau et al., 2008; Hertwig & Erev, 2009; Hills & Hertwig, 2010) make the "information acquisition" possibility less likely. If exploration was used to obtain information without concerns about identifying the option that provides the maximum rewards, the number of samples would be larger. However, the search process during sampling is, in fact, costly (Fiedler & Kareev, 2006; Kareev, 2000; Hau et al., 2008), and there might be some advantages to fewer samples. Studies have shown that fewer samples may render the choice simpler and surprisingly good (Hertwig & Pleskac, 2008, 2010), because fewer samples lead to larger initial differences between two options being considered, compared to the differences given by their objective probabilities (i.e., the "amplification effect"). Furthermore, this differentiation between the two options with small samples may ease the choice process and may lead to choices that, although are not optimal, are good enough (Hertwig & Pleskac, 2008, 2010).

Hau et al. (2008) demonstrated that people consider and account for perceived costs during sampling. Their studies show that sampling is costly in terms of opportunity costs. For example, sampling might take time during which people cannot pursue other activities. Furthermore, they demonstrated that people consider the magnitude of outcomes when deciding whether or not to continue sampling the options: When the values of the outcomes were increased (resulting in higher opportunity costs from not choosing the option with the higher expected value), the sample size doubled (a median of 33) compared to the same problems in Hertwig et al. (2004). Thus, the amount of search does depend on the value of the outcomes involved. Furthermore, Gonzalez and Dutt (2012) found a pattern of decreased exploration with increased sampling in Hau et al.'s (2008) data. This pattern occurred at the average and individual participant levels. They demonstrated that the patterns of decreased exploration during sampling occur regardless of the sample length and that the frequency of sampling-H was indicative of the final choice. These patterns of reduced exploration in a sampling paradigm are very similar to those found in consequential choice paradigms, leading the authors (Gonzalez & Dutt, 2011, 2012) to suggest the presence of an exploration-exploitation tradeoff during free sampling similar to that found in consequential choice (Biele, Erev, & Eyal, 2009; Camilleri & Newell, 2011; Gonzalez & Dutt, 2011; Mehlhorn et al., 2014).

The studies reviewed above provide support for the idea that people explore options during free sampling in a way that the process is led by the economic value of the options rather than by the pure epistemic value. That is, the search process may serve to discover and pursue the maximizing option, and then the patterns of decreased exploration rate should be inversely related to patterns of increased sampling-H rate over more samples.

Search Strategies: How do people explore in a free sampling binary task

Hills and Hertwig (2010) used data from experiments in the sampling paradigm to investigate the strategies that humans may use during sampling. They used the alternation rate between the two options to investigate two prominent sampling strategies: *Piecewise*, where options are explored very rapidly and participants alternate back-

and-forth between them in a zigzag manner (see Figure 2, left panel); and *comprehensive*, where participants explore one option more deeply before making a switch to the other and switching back-and-forth between the options is less frequent (see Figure 2, right panel).

Hills and Hertwig (2010) discovered that although our search may reveal the same information, the strategies we use during sampling influence the subsequent choice. For example, the piecewise strategy more often resulted in the *underweighting of rare outcomes* (e.g., $0 in the example shown in Figure 2) compared to the comprehensive strategy. They also found that the piecewise strategy resulted in less consistency (agreement) between the predictions from sampling behavior and the consequential choice.

However, Hills and Hertwig (2010) left one important question unanswered (page 5): "Why is the way people search indicative of the final decisions they make?" We expect the dynamics of exploration and exploitation and the inverse relationship between exploration rate and sampling-H rate to be the answer to this question. Gonzalez and Dutt (2011, 2012) suggested that the main features of decisions from experience can be captured with the hypothesis that people tend to select the option that led to the best value in similar situations in the past. A formalization of this underlying process is provided by Instance-Based Learning Theory (IBLT) (Gonzalez, Lerch, & Lebiere, 2003). In essence, IBLT proposed that decisions are made by retrieving experiences from past similar situations and selecting the option that led to the best outcomes. In agreement with other Instance-based theories of learning (Dienes & Fahey, 1995; Logan, 1988) and reinforcement-learning processes (Erev & Roth, 1998), IBLT proposes that depending on the consistency or variability of environmental conditions, there would be a gradual transition from exploration to exploitation of options that have provided the best outcome based on experience.

The main choice rule in IBLT is to select the option with the maximum experienced value (called *Blending*) (Gonzalez & Dutt, 2011; Lejarraga, Dutt, & Gonzalez, 2012). When the blended value of one option (A) is higher than the blended value of the other option (B), choose A, otherwise choose B. In the example of Figure 1, there are three instances, each corresponding to each possible outcome: (A, 4), (A, 0), and (B, 3). Each option (A and B) has a blended value calculated as the sum of each experienced outcome based on the *cognitive* probability (probability of recalling that outcome from memory). The cognitive probability of an instance is determined by several memory factors including the frequency and recency (memory decay) (these are components of an *Activation* mechanism obtained from the ACT-R theory of cognition; Anderson & Lebiere, 1998). If an outcome has been experienced more often and more recently, that instance would have higher activation, which increases its probability of retrieval (see formalization of these mechanisms in Gonzalez & Dutt, 2011; and Lejarraga et al., 2012). IBLT's mechanisms predict a process in which there is a gradual transition from more exploration of the available options towards exploitation of options that have resulted in the best outcomes through experience (Gonzalez et al., 2003).

Given the possibility that people would strategize about how to explore the options, selecting a preferred strategy over other strategies (Hills & Hertwig, 2010), we question if the same or different dynamics may emerge during sampling when piecewise or comprehensive strategies are used. Note that the piecewise and comprehensive strategies are

"idealized"; that is, they are two extremes of a continuum of alternation or exploration processes (Hills & Hertwig, 2010). In what follows, we analyze the dynamics of exploration and sampling-H overall and under piecewise and comprehensive strategies, by relying on a large data set of the sampling paradigm that is publicly available (Erev et al., 2010).

Method

Two data sets from the TPT's sampling competition were put together: an estimation set (60 problems) and a competition set (60 new problems derived using the same algorithm as the estimation set), which are both available online (Erev et al., 2010). All problems involved sampling between two unlabeled buttons, one associated with a safe option that offered a medium (M) outcome with certainty and the other associated with a risky option that offered a high (H) outcome with some probability (pH) and a low (L) outcome with the complementary probability (1-pH) (see Erev et al., 2010 regarding the problem generation algorithm and data collection methods).

In each of the estimation and competition sets, 40 participants were randomly assigned into two groups of 20 participants each, and each group completed 30 of the 60 problems in a random order.[1] Participants were allowed to sample the options freely as long as they wanted and in their desired order before making a consequential final choice. Although participants could sample freely; however, the median sample size across the two options was small (= 9 samples). Using the same assumption as in Hills and Hertwig (2010), we only considered those problems where participants saw all the outcomes for both of the options[2], obtaining a data set with 74 participants, 120 problems, encompassing 18,113 sampling decisions, and 988 observations (observations is a unique combination of participant, problem, and set that is used as the unit of our analysis).

We calculated the A-rate as done by Gonzalez and Dutt (2011, 2012) and Hills & Hertwig (2010). For each observation starting in the second sample, we coded whether the participant switched the choice (=1) or not (=0) from the previous sample (the very first sample was marked as a missing value as there was no sample preceding it). Then, the alternation rate was defined as the average at each sample computed across observations. To calculate the Sampling-H rate we first identified the option with the high expected value in each problem. Based on the definition of those problems, in 63 problems the high expected
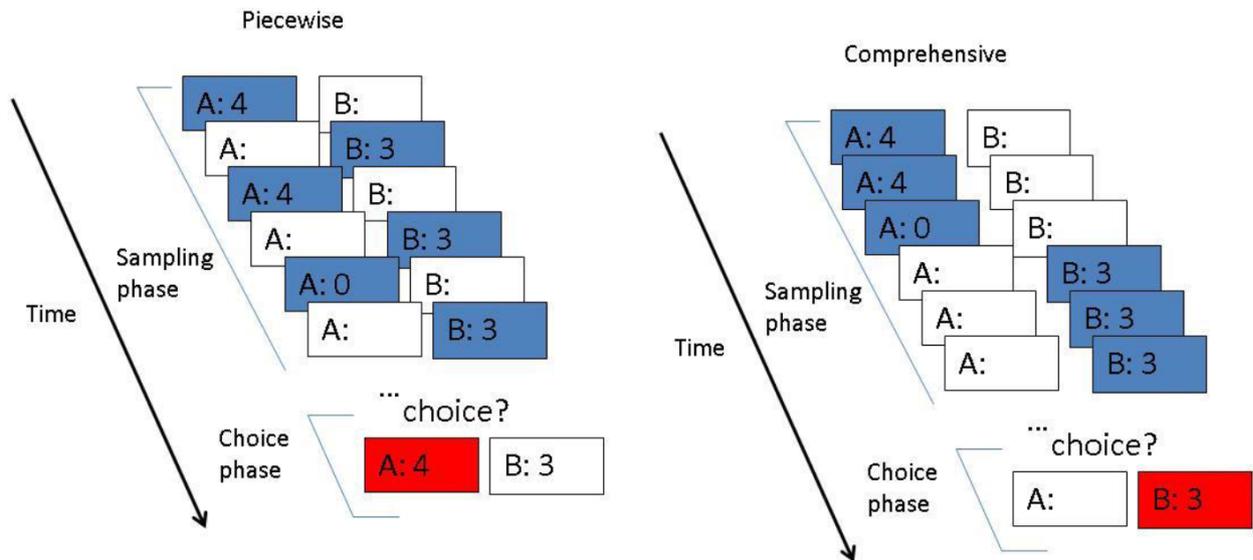

Figure 2. Examples of piecewise (left panel) and comprehensive strategies (right panel). When a person decides to stop the search they are asked to make a consequential choice. In this example, risky option A: represents a .8 chance to get $4 and .2 chance to get $0 and safe option B: gets $3 for sure. Choices are influenced by the strategy of search according to Hills & Hertwig (2010).

value option was the safe option, and in 57 problems it was the risky option. Then, we checked whether the option sampled by a participant was the high expected value option, and coded this as 1; otherwise, the choice was coded as 0. We then aggregated high choices across all participants and problems for different samples and defined the Sampling-H rate per sample.

Results

Figure 3 shows the overall A-rate and Sampling-H rate across samples including all observations in the data set. This figure shows all sample trials up to the point in which there were at least two observations left in the data set (sample number 133). The figure shows a gradual increase of the Sampling-H and in parallel, a gradual decrease of the A-rate with increased sample trials. Given that people rely on small samples (Hills & Hertwig, 2010; Hau et al., 2008) the number of observations decreases rapidly with increased samples. This explains the noisy averages as sample sizes increase, given that they involve fewer participants (Gonzalez & Dutt, 2011, 2012; Hills & Hertwig, 2012). Using the median sample size of the overall data set (Median = 10) with a Cochran's Q test, we found a significant difference in the A-rate across the first 10 samples, $\chi^2(8)$ = 81.66, $p < .001$) and a significant difference in the Sampling-H rate across the first 10 samples, $\chi^2(8)$ = 39.84, $p < .001$. A pairwise comparison revealed a decrease in the A-rate from .40 in sample #2 to .23 in sample #10, a 43% drop ($Z = -7.14$, $p < .001$); and an increase of 13% in the Sampling-H rate from .52 in sample #2 to .60 in sample #10, ($Z = -1.98$, $p < .05$). This result suggests that across samples, participants explore between the two buttons less, while increasingly selecting the option with the higher expected value. In fact, the Sampling-H rate was significantly and negatively correlated to the A-rate, $r_s = -.48$, $p < .01$.
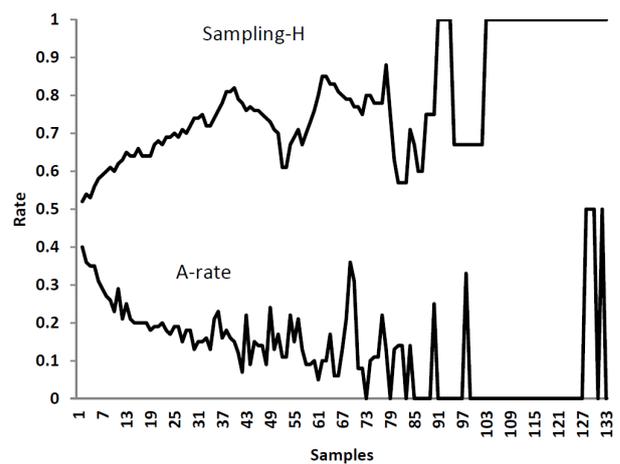

Figure 3. The average Sampling-H rate and A-rate across samples.

Sampling-H and A-rate for piecewise and comprehensive search strategies

To analyze behavior for different sampling strategies, we first analyzed the distribution of the alternation rate between the two options (see Figure 4) and followed Hills and Hertwig's (2010) procedure of classifying participants according to their A-rate. The A-rate in the TPT data set varied widely, from a minimum of 0.07 to a maximum of 1.0. The median in the TPT data set was higher (.27) (shown by the dotted line in Figure 4) than in Hills and Hertwig's (2010) data (.16), but the distribution was similarly bimodal with peaks in the 0.15-0.20 and 0.45-0.50 A-rate intervals. Accordingly, all participants with an A-rate less than 0.27 were categorized as following a comprehensive strategy; whereas, all participants with an A-rate

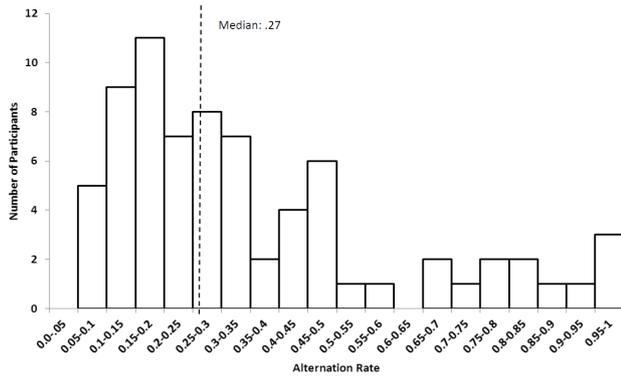above 0.27 were categorized as following the piecewise strategy.

Figure 4. Histogram of participants' A-rate (averaged across all problem played by a participant in a set). The A-rate in a problem for a participant is expressed as the ratio of observed switches and the maximum number of allowable switches (n-1, where n are the number of samples) in a problem. The dotted line represents the median value = 0.27.

Figure 5 shows the A-rate and Sampling-H rate for piecewise (left panel) and comprehensive (right panel) strategies. The maximum trial in which there were more than one observation left in the data set was 72 for the piecewise strategy and 133 for the comprehensive strategy. That is, people who alternated more often tended to take fewer samples than people who alternated less often. The median sample size for the piecewise group was 8, while the median sample size for the comprehensive group was 12. This result supports similar observations by Hills and Hertwig (2010), and also Rakow, Demes, and Newell (2008).

A general observation of these patterns indicates a decrease in A-rate over increased trials and an increase in the Sampling-H rate regardless of the sampling strategy. For participants following a piecewise strategy, there was a significant decrease in the A-rate ($\chi^2(6) = 82.17$, $p < .001$) and a significant increase in the Sampling-H rate ($\chi^2(6) = 21.69$, $p < .01$) over the first 8 samples. For participants following a comprehensive strategy, there was a significant decrease in the A-rate ($\chi^2(10) = 42.44$, $p < .001$), but not a significant increase in the Sampling-H rate across 12 samples ($\chi^2(10) = 2.29$, $p = .99$). Although the trend of Sampling-H rate increases over time on average, this result may be due to the different orders in which participants may sample one or the other option (see the discussion of results). For example, it is possible that some participants start by exploring the high option and then move to the low expected value option, and others do the reverse order in the comprehensive strategy. Although these clean patterns of exploration are only idealistic in the comprehensive strategy, what matters in this research is that for both comprehensive and piecewise strategies the Sampling-H rate was significantly and negatively correlated to the A-rate (comprehensive: $r_s = -.35$, $p < .01$; piecewise: $r_s = -.24$, $p < .05$).

Consistency between sampling and final choice

Figure 6 reports the proportion of total agreement between predicted final choice based upon Sampling-H rate during sampling and participant's actual final choice. For this analysis, we classified participants based upon the median Sampling-H rate (similar to how participants were classified as following the piecewise and comprehensive strategies). The median Sampling-H rate during sampling was 0.50. Observations below this rate were classified as infrequent Sampling-H, and those at or above 0.5 were classified as frequent Sampling-H. Among the frequent and infrequent Sampling-H, we also identified those observations that followed the piecewise strategy (median alternation rate > 0.27) and the comprehensive strategy (median alternation rate < 0.27). Within each of the four combinations of sampling strategy and Sampling-H rate, we calculated the average of the outcomes obtained during sampling in each option. As per IBLT, the option with the highest average would be the one that is predicted to be chosen at the final choice. We matched the predicted final choice based upon the highest average with the actual final choice. Next, we calculated the proportion of agreement by averaging such matches across all observations in each of the four combinations.

As observed in Figure 6, regardless of the alternation strategy and frequency of Sampling-H, there is a high consistency ($> 50\%$) between predicted final choices at the end of sampling and the actual final choice made by participants. The consequential choice after sampling was equally predicted for piecewise and comprehensive strategies, and for both infrequent ($Z = -0.913$, $p = .38$) and frequent Sampling-H participants ($Z = -0.642$, $p = .53$).
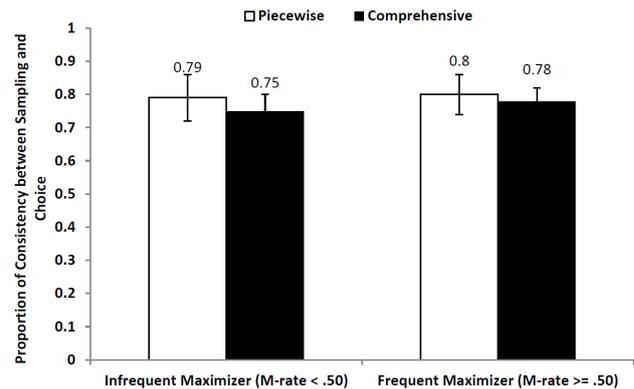

Figure 6. Consistency between sampling behavior and consequential choice for frequent and infrequent Sampling-H participants following the piecewise and comprehensive strategies.

Discussion

Our results clarify the relationship between the rate of exploration and the tendency to explore the option with high value during free sampling. We find a decrease in exploration rate and an increase in the rate
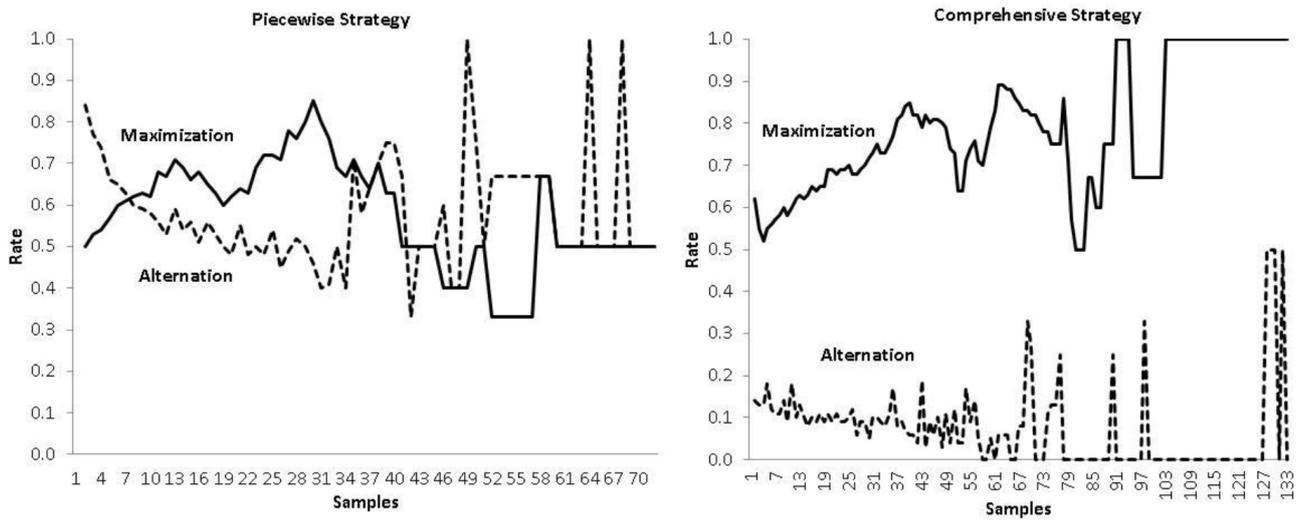
Figure 5. The A-rate and Sampling-H rate across samples for the piecewise and comprehensive strategies.

of sampling of the option of high value. Our results show significant inverse dynamics between the A-rate and the Sampling-H rate in binary-choice problems. Furthermore, these negatively correlated dynamics appear regardless of the search strategies people might adopt during sampling. Finally, we also show that the final consequential choice can be accurately predicted by the frequency of selection of the high option during sampling.

These results are important as they provide support to an initial suggestion (Gonzalez & Dutt, 2011, 2012) that a decrease in the exploration rate during sampling is related to the process of discovering the option that maximizes expected value. Our results indicate that free sampling is not a simple random process where participants explore the options with the goal of informing their future decisions. Rather, we show that learning during free sampling is a gradual discovery of the best option while reducing the exploration effort with more samples. As suggested by theoretical accounts of decisions from experience, participants seem to gradually move from a process of exploration to the exploitation of the best option, and they end up choosing the option with agrees to this patterns of sampling (Gonzalez et al., 2003).

As in Hills and Hertwig's findings (2010) we also identified two idealized search strategies: piecewise and comprehensive. However, regardless of which strategy was used, the search process seemed to serve the same purpose as demonstrated by similar increase in the sample rate from the high value option, while this process is inversely related to a gradual decrease in exploration. These phenomena are explained by IBLT's learning process which suggests that choice is led by a dynamic formation of the value of the options through experience (*Blending*). The process of discovering the most valuable option starts with more exploration (reflected in higher alternation between the two options in binary choice), but as the better option becomes evident through experience, the amount of exploration is reduced. We find that regardless of the

exploratory strategy, people exhibit similar dynamics of decreased exploration and increased selection of the high value option over time. For example, with a piecewise strategy and using the example of Figure 1, the activation of the outcome for the safe option ($3, in the example) will be high given the frequency of selection of this option; while the activation of the outcomes for the risky option ($4 and $0) will vary according to the frequency with which these outcomes are observed. The $4 instance is experienced more often and could result in a higher activation than the $0 instance, which is a rare event (the activation equation has some stochastic noise, see Gonzalez & Dutt, 2011). When $0 and $4 are combined through Blending, it is expected that the blended value of the risky option would more often be slightly higher than the one of the safe option and as a result, the risky option is expected to be chosen increasingly over the safe option, resulting in a gradual reduction of alternation between the two options where the risky option is often selected (here, risky option is also the high value option and the selection of this option increases the Sampling-H rate).

With a comprehensive strategy and using the example of Figure 1, the activation of the three instances would greatly depend on the order in which the options are selected and on the number of times that an option is consecutively selected. If the risky option is selected first (as in the example of Figure 2, right panel) and then a switch is made to the safe option, the activation of the outcomes for the risky option ($4 and $0) would decay during the longer exploration of the safe option. This order would increase the chances of choosing the safe option more often than the risky (low Sampling-H rate) and decrease alternation between the two options. The reverse order of exploration would predict an increased chance of choosing the risky over the safe option, resulting in higher Sampling-H rate.

In conclusion, this research contributes towards understanding the relationships between exploration-exploitation processes during free sampling, their dynamics, and their consequences for choice. Our results

indicate that regardless of explicit search strategies, a decrease in exploration is observed in parallel to an increase in the selection of the high value option. However, a general conclusion regarding these phenomena is expected to depend on the dynamics of probabilities and outcomes of the environment over the course of free sampling. In highly dynamic environments, the diversity of options would make it more challenging for humans to discriminate among familiar classes of objects and more exploration would be required. Although decisions might become increasingly similar with task practice, higher Sampling-H rates might be slower in dynamic and diverse environments. Understanding and predicting the rate at which exploration decreases and Sampling-H rate increases has important implications for training and learning from experience. Presumably, one could strategically manipulate the speed of these transitions by introducing surprising outcomes during sampling, which may keep people interested in alternating between options (thus, inviting increased exploration and delay exploitation). Furthermore, another likely way of influencing the speed of of these transitions may be via introducing more options. When there are more than 2-options to choose from, it is likely that transitions will be delayed compared to when confronted with just 2-options. That is because more options in the choice set would likely make it difficult for people to find the options of high value. Some of these ideas form the immediate next steps for us to investigate in the near future.

Declaration of conflicting interests: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

Author contributions: The authors contributed equally to this work.

Supplementary material: No supplementary material available.

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought.* Hillsdale, NJ: Erlbaum.

Biele, G., Erev, I., & Eyal, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology, 53*(3), 155-167. doi:10.1016/j.jmp.2008.05.006

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review, 18*(2), 377-384. doi:10.3758/s13423-010-0040-2

Dienes, Z., & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(4), 848-862. doi:10.1037/0278-7393.21.4.848

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S., Hau, R. Hertwig, R., Stewart, T., West, R., & Lebiere, C. (2010). A choice prediction competition for choices from experience and from description. *Journal of Behavioral Decision Making, 23*, 15-47. doi:10.1002/bdm.683

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review, 88*(4), 848-881. doi:10.1002/bdm.683

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107*(4), 659-676. doi:10.1037/0033-295X.107.4.659

Fiedler, K., & Kareev, Y. (2006). Does decision quality (always) increase with the size of information samples? Some vicissitudes in applying the law of large numbers. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32*(4), 883-903. doi:10.1037/0278-7393.32.4.883

Gonzalez, C., & Dutt, V. (2011). Instance-Based Learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118*(4), 523-551. doi:10.1037/a0024558

Gonzalez, C., & Dutt, V. (2012). Refuting data aggregation arguments and how the IBL model stands criticism: A reply to Hills and Hertwig (2012). *Psychological Review, 119*(4), 893-898. doi:10.1037/a0029445

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27*(4), 591-635. doi:10.1016/S0364-0213(03)00031-4

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. doi:10.1002/bdm.598

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15*(8), 534-539. doi:10.1111/j.0956-7976.2004.00715.x

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517-523. doi:10.1016/j.tics.2009.09.004

Hertwig, R., & Pleskac, T. J. (2008). The game of life: How small samples render choice simpler. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 209-235). Oxford, UK: Oxford University Press.

Hertwig, R., & Pleskac, T. J. (2010). Decision from expe-

rience: Why small samples? *Cognition, 115*(2), 225-237. doi:10.1111/j.0956-7976.2004.00715.x

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science, 21*(12), 1787-1792. doi:10.1177/0956797610387443

Hills, T. T., & Hertwig, R. (2012). Two distinct exploratory behaviors in decisions from experience: Comment on Gonzalez and Dutt (2011). *Psychological Review, 119*(4), 888-892. doi:10.1037/a0028004

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review, 107*(2), 397-402. doi:10.1037/0033-295X.107.2.397

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making, 25*(2), 143-153. doi:10.1002/bdm.722

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492-527. doi:10.1037/0033-295X.95.4.492

Mehlhorn, K., Ben-Asher, N., Dutt, V., & Gonzalez, C. (2014). Observed variability and values matter: Towards a better understanding of information search and decisions from experience. *Journal of Behavioral Decision Making, 27*(4), 328-339. doi:10.1002/bdm.1809

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643-675. doi:10.1037/0033-295X.106.4.643

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes, 106*(2), 168-179. doi:10.1016/j.obhdp.2008.02.001

Rakow, T. & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making, 23*(1), 1-14. doi:10.1002/bdm.681

Teodorescu, K., & Erev, I. (2014). On the decision to explore new alternatives: The coexistence of under- and over-exploration. *Journal of Behavioral Decision Making, 27*(2), 109-123. doi:10.1002/bdm.1785

Todd, P. M., Penke, L., Fasolo, B., & Lenton, A. P. (2007). Different cognitive processes underlie human mate choices and mate preferences. *Proceedings of the National Academy of Sciences, 104*(38), 15011. doi:10.1073/pnas.0705290104

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science, 23*(7), 698–703. doi:10.1177/0956797612443968