A preference for the unpredictable over the informative during self-directed learning

Doug Markant (markant@mpib-berlin.mpg.de)
Center for Adaptive Rationality
Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, Germany

Todd M. Gureckis (todd.gureckis@nyu.edu)
New York University
Department of Psychology
6 Washington Place, New York, NY 10003 USA

Abstract

The potential information gained from asking a question and one's uncertainty about the answer to that question are not always the same. For example, given a coin that one believes to be fair, the uncertainty a person has about the outcome of flipping that coin is high, but either outcome is unlikely to make them believe that the coin is biased (i.e., the "information gain" of that observation is low). In the present paper we show that people use a simple form of *predictive uncertainty* to guide their information sampling decisions, a strategy which is often equivalent to maximizing information gain, but is less efficient in environments where potential queries vary in their reliability. We conclude that a potentially powerful driver of human information gathering may be the inability to predict what will happen as a result of an action or query.

Keywords: self-directed learning, active learning, information search

How do people make decisions to gather information? One common approach for answering this question is to compare information gathering behavior of humans to normative theories that objectively measure the information contained in particular queries. For example, *information gain* (IG) measures the usefulness of a potential query based on how much it is expected to reduce uncertainty. IG has been shown to account for sampling decisions in a number of search problems (Nelson, 2005), with recent work suggesting that it can also predict eye movements during visual search, including when searching for a target embedded in noise (Najemnik & Geisler, 2005), disambiguating the identity of a shape (Renninger et al., 2007), and when learning the relevance of features for categorization (Nelson & Cottrell, 2007).

However, it is important to note that normative models (even when they fit well) do not necessarily reveal people's underlying decision making process. For example, in many situations, simpler decision making strategies are consistent with IG, showing that judgments about the rationality of any single preference during information search can only be made with respect to the structure of the learning environment (Klayman & Ha, 1987). For example, people often use *positive testing* (PT), a preference for positive instances or strongly predicted features. Although PT has often been portrayed as a form of confirmation bias, it has been shown to mimic IG under common kinds of problems where positive evidence is rare but highly informative (Klayman & Ha, 1987; Navarro & Perfors, 2011).

Similarly, under some conditions IG is equivalent to a simple preference for observations whose outcomes are difficult to predict, which is captured by a sampling model we refer to as *label entropy* (LE, where a "label" is the feedback received from an observation). Specifically, when hypotheses are deterministic (i.e., they predict the outcome of a given observation with probability 0 or 1), IG and LE make identical predictions as to the value of different queries (Markant et al., in revision). The distinction between IG and LE can be understood intuitively in terms of a possibly biased coin toss. If one is confident that the coin is fair, then a single flip (heads or tails) is unlikely to change that belief. At the same time, it is difficult to predict what the outcome of the coin flip will be. Whereas evaluating the IG of a potential query implies a consideration of how a query might change one's beliefs, LE only requires a person to evaluate how confidently they can predict its outcome in light of what they already know.

In prior experiments using the task described in this paper, we found that people make sampling decisions that are consistent with IG (Gureckis & Markant, 2009; Markant & Gureckis, 2012), but in those tasks LE made identical predictions owing to the use of deterministic hypotheses. The goal of the current experiment was to differentiate between LE and IG by decoupling the amount of information conveyed by a query from the uncertainty about its outcome. One way to achieve this is for queries to vary in their *reliability*, that is, the probability of receiving accurate feedback about their outcome.

Varying the reliability across potential queries allows us to differentiate possible models of human information sampling. Given an information source for which reliability is low, a learner should be highly uncertain about the outcome of querying it, while the expected information conveyed should be low. If people rely solely on their uncertainty about the outcome of a particular observation in order to judge its value, they should choose in accordance with LE. Alternatively, if they evaluate how informative a query is expected to be, taking into account any chance of incorrect feedback, they

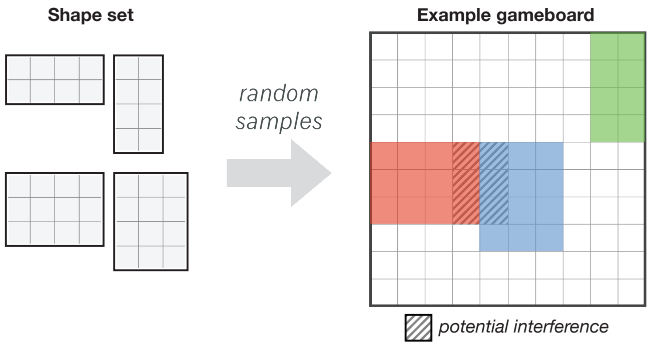

Figure 1: Set of shapes used to generate gameboards (left) and an example gameboard (right). "Interference" could occur in locations adjacent to a different shape, leading to incorrect feedback about the color in those locations.

should behave in accordance with IG. This approach allows us to both test the generalizability of IG as a computational-level account of information search, and to understand the extent to which deviations from optimal sampling arise from simple uncertainty-driven decision making.

In the current study we tested whether people made adaptive sampling decisions in a sequential, spatial learning task, using an ideal observer to model how uncertainty about the target concept changes over the course of learning (for a similar approach, see Nelson & Cottrell, 2007; Najemnik & Geisler, 2005). Our goals were to evaluate whether people could optimally adjust how they search for information given the constraints of the task, and to characterize any deviations from a normative standard (IG) using two alternative accounts of how people make sampling decisions: *positive testing* (PT) and *uncertainty sampling* (LE). Our results show that people do not make decisions most consistent with IG, but instead rely on a simpler form of predictive uncertainty in order to select between information sources.

The search task

Like the board game *Battleship*, in the "Ship Surveillance Game" the player is presented with a 10x10 grid containing three hidden shapes (see Figure 1). There are two phases in each game: a *sampling phase* and a *test phase*. In the sampling phase, the participant learns about the hidden shapes by choosing squares in the grid to uncover, revealing either part of a hidden shape or an empty square. When they think they know the the location and form of the shapes, they can choose to stop sampling and enter the test phase, at which point they "paint" any remaining squares they believe belong to one of the shapes in the appropriate color.

In the game there are monetary costs for making observations and for making errors during the test phase. The specific costs we used (with painting errors more costly than observations) promote efficient information search in two ways. First, the observation cost discourages sampling in locations whose contents can be inferred from evidence that has already been uncovered. Second, the cost of errors encourages continued sampling while there is still uncertainty about the hidden rectangles. The goal of the player is thus to minimize their penalties by learning the true configuration of shapes in the fewest observations possible.

Over the course of multiple games, each participant is faced with unique arrangements of targets and makes highly variable sequences of observations. On any given trial, the value of querying a location is dynamically related to the set of shape configurations that are still plausible. In order to understand this diverse sequence of choices, we used a Bayesian ideal observer to represent the uncertainty about the true game board on each trial given all previous observations. Based on this model of uncertainty in the task, we then compared the predictions of three sampling models to participants' choices.

Ideal observer model

In each game, a player must learn a hidden gameboard corresponding to a single hypothesis h in the universe of legal gameboards, H. At the beginning of each trial t, the player has seen previous observations, $\mathcal{D} = \{(x_1,l_1),(x_2,l_2),...,(x_{t-1},l_{t-1})\}$, where x denotes a location in the grid and l the feedback received (when $l=0$, that location is empty; if $l=1$ if contains part of the first shape, and so on). The posterior probability distribution over the hypothesis space H is given by Bayes' Rule:

$$p(h|\mathcal{D}) = \frac{p(\mathcal{D}|h)p(h)}{\sum_{h' \in H} p(\mathcal{D}|h')p(h')} \quad (1)$$

where the prior $p(h)$ is a uniform distribution over the hypothesis space (the likelihood $p(\mathcal{D}|h)$ is defined below). The posterior distribution is then used to predict the probability of a new query x resulting in the outcome l, given by:

$$p(l|x,\mathcal{D}) = \sum_{h \in H} p(l|x,h)p(h|\mathcal{D}) \quad (2)$$

The state of the ideal observer reflects any remaining uncertainty about both the true gameboard and the outcome of any remaining queries. The goal of a sampling model is a quantitative measure that, for each potential observation x, transforms this uncertainty into a predicted value $V(x)$ of sampling that location.

On each sampling trial t there is a set of observations X_t available to be sampled, from which the learner makes a query $x \in X_t$. Each model assigns a predicted value $V(x)$ to each observation x (as defined below). These values are converted into choice probabilities using the softmax function:

$$p(x) = \frac{e^{V(x)/\beta}}{\sum_{y \in X_t} e^{V(y)/\beta}} \quad (3)$$

The temperature β is fit to each participant by maximizing the summed log-likelihood of their sampling decisions from all games played. The log-likelihood corresponding to the best-fit value of β is used to determine which model provides the best account of each participant's decisions.

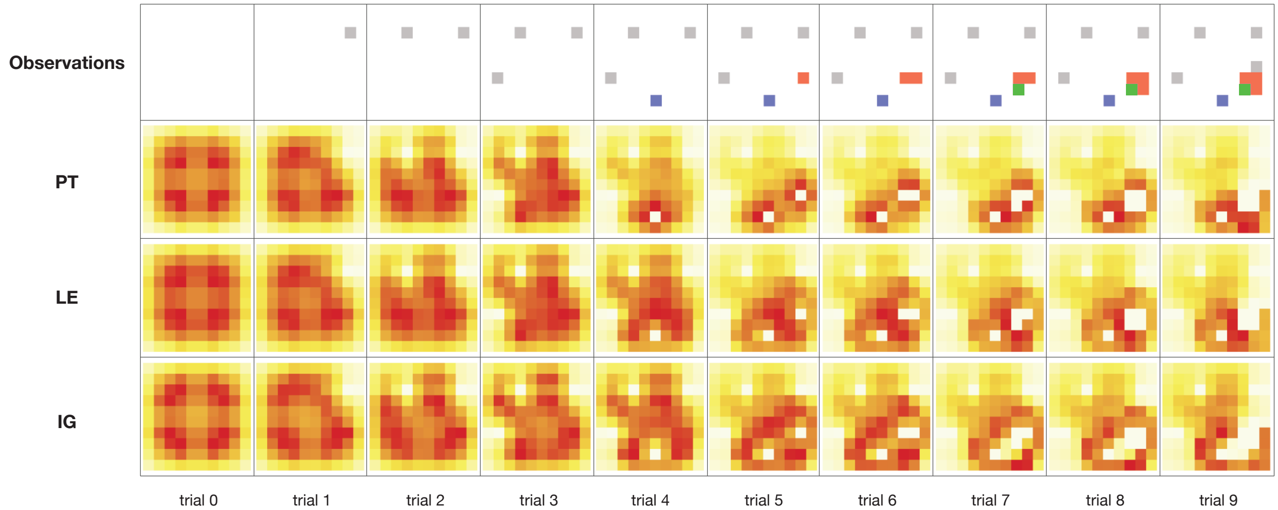

Figure 2: A participant's first 9 selections (top row) and the predicted value of new observations $V(x)$ for three sampling models. The model predictions show the relative value of selecting different queries, with red locations more likely to be chosen.

Information gain (IG) Information gain predicts that a query x is valued according to the expected reduction in uncertainty about the true hypothesis, averaged across all possible outcomes of a query:

$$V_{IG}(x) = \sum_{n=0}^{3} p(l_n|x,\mathcal{D})\,[I(D) - I((x,l_n),D)] \qquad (4)$$

$I(D)$ is the Shannon entropy over the posterior distribution given a set of observations $\mathcal{D}$:

$$I(\mathcal{D}) = -\sum_{h\in H} p(h|\mathcal{D})\, log\; p(h|\mathcal{D}) \qquad (5)$$

Label entropy (LE) The label entropy model predicts a preference for those features whose outcome is most uncertain. Given the predictive distribution for a new query x, the value is defined as the Shannon entropy across possible outcomes l of that query:

$$V_{LE}(x) = -\sum_{n=0}^{3} p(l_n|x,\mathcal{D})\; log\; p(l_n|x,\mathcal{D}) \qquad (6)$$

Queries that are strongly predicted to result in a single label will not be favored under this model. Instead, those locations about which the outcome is highly uncertain will be preferred.

Positive testing (PT) Positive testing indicates a preference for queries that are expected to result in positive evidence of a target concept (i.e., a "hit"). In our task we do not ask participants to identify their current hypotheses about the gameboard, and as a result can't evaluate whether their observations are positive tests. Instead, we model positive testing as the probability of getting positive feedback (i.e., any non-zero label) based on the state of the ideal observer:

$$V_{PT}(x) = \sum_{n=1}^{3} p(l_n|x,\mathcal{D}) \qquad (7)$$

Thus, the PT model will assign a high value to an observation that has not yet been uncovered but is predicted to be part of a target, while assigning zero value to any observations that are predicted to be empty locations.

Distinguishing between alternative sampling models

As we noted above, under some conditions PT is consistent with IG, and this was also true during the early stage of our task. Figure 2 shows the first 10 trials of a game, along with the predicted value of the three models (with red indicating a stronger predicted preference for querying that location). On the first few trials of the game, the predictions of all three models are nearly identical. After some observations are accumulated, however, PT diverges strongly from the other two. In particular, it predicts sampling locations that are known to be part of a target, even though the outcome can be inferred and there is no information to be gained (since this particular kind of choice is at odds with the cost structure of the game, we refer to it as a *confirmatory error* in our analysis).

In order to distinguish between IG and LE we introduced "interference" to the task such that when any two rectangles were adjacent in the true gameboard, any squares that touched a different target had a 50% chance of producing an incorrect label (with all incorrect labels equally likely, including the "empty" label l_0). As a result, observations varied in their reliability depending on the likelihood of interference occurring. For example, on trial 5 in Figure 2, interference is likely to occur in the region between the blue and red squares, leading to a lower predicted value according to IG than LE.

The likelihood function was defined in the following way. For a given hypothesis h, queries were divided into two sets: those that belonged to a target shape and were adjacent to a different target, X_{adj}, and any other observations, X_{nonadj}. For $x \in X_{nonadj}$, the likelihood $p(l|x,h)$ was deterministic:

$$p(l|x,h) = \begin{cases} 1 & \text{if } h(x) = l \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

where $h(x)$ is the true label at location x. For $x \in X_{adj}$, however, the likelihood incorporated the noisy feedback:

$$p(l|x,h) = \begin{cases} 1/2 & \text{if } h(x) = l \\ 1/6 & \text{otherwise} \end{cases} \quad (9)$$

Just as the value of different queries is linked to the set of hypotheses remaining, the chance that a location in the grid would produce noisy feedback depended on the distribution of plausible targets. Rather than learning that any particular location was less reliable than others, accounting for the possibility of incorrect feedback depended on the learner's current knowledge of potential configurations of targets.

In our experiment we cast the probabilistic nature of the feedback as arising from noise, rather than probabilistic hypotheses. This led to a distinction between a learner's uncertainty about the feedback received during sampling and their uncertainty about the prediction they would make during the test. For example, based on previous observations a participant might be perfectly confident that a location belonged to the red target, while also knowing that interference could occur in the same location (resulting in uncertainty about the outcome of querying it). Since either type of uncertainty could potentially guide sampling decisions, we compared two sets of PT and LE models: one set that accounted for noisy feedback that could occur during sampling (PT and LE), and one set that ignored noisy feedback in order to reflect uncertainty about the underlying label (PT_{det} and LE_{det}).

Three models (PT, LE, and IG) used the noisy likelihood function (Eqns. 8 and 9 as described above) to predict the value of new observations. For the PT and LE models, this implies that people use their uncertainty about the outcome of the observation during the sampling phase to judge its value. For the PT model, a high probability of interference implies a lower probability of hitting a target (and a weaker preference). The reverse occurs for the LE model: locations with a high probability of interference are associated with higher uncertainty about the outcome of the observation, implying a stronger preference.

The remaining two models (PT_{det} and LE_{det}) predicted that people evaluate queries based on their ability to predict the true label during test (which is not affected by the interference that occurs during sampling). For these models we used the deterministic likelihood function in Equation 8 when computing the label predictions for all locations. The PT_{det} model thus valued any observations that were predicted to be a part of a target, regardless of whether or not interference might occur. Similarly, the LE_{det} model valued an observation about which the person was less confident about the true label, irrespective of whether querying that location was likely to result in an incorrect label. Figure 3 illustrates the divergence between these models for an example state in the game.

Experiment

Participants

Fifty participants were recruited using Amazon Mechanical Turk. Each person was paid $2 for their participation and awarded a bonus depending on their performance. Seven participants were excluded from the analysis because they skipped at least one game (i.e., by making zero observations).

Materials

Gameboards were generated using a set of four rectangle shapes of varying size (2x4, 4x2, 3x4, or 4x3). Each gameboard was created by randomly sampling three rectangles from this set (with replacement) and randomly placing them on a 10x10 grid (see Figure 1). Each participant played 30 games. In 2/3 of games the underlying gameboard contained adjacent shapes, while the remaining 1/3 of games did not.

Procedure

The task was described as a "Ship Surveillance" game in which the player located ships within a patch of ocean by searching for the presence of radio signals. If any two ships were adjacent to each other there was "interference" such that the signal in any location adjacent to a different ship was noisy, leading to incorrect feedback 50% of the time (see Figure 1). If interference occurred, feedback was randomly selected from the set of incorrect labels with equal probability.

Each game began with a possible bonus of $5. During the sampling phase, a participant made an observation by clicking on a square, causing it to change color according to its membership in one of the shapes. After a square was uncovered, it remained visible for the remainder of the game. Each observation decreased the amount of the potential bonus by $0.20. Participants were instructed to stop sampling when they felt they had learned the hidden shapes, at which point they entered the test phase and filled in any remaining squares they believed were part of a hidden shape. Each error they committed (either by failing to fill in part of a target or filling in a square incorrectly) decreased their potential bonus by $2.00 (this high penalty was used to discourage participants from ending the sampling phase before they were confident about the locations of the targets). Lastly, the true gameboard was shown (with 'x's marking squares that were painted incorrectly) along with the final bonus, including the amount of penalties that were attributable to the number of observations made and how many were the result of painting errors.

The instructions emphasized that the key to good performance was to learn the true gameboard in the fewest number of observations possible. Following the instructions, participants were shown 50 random gameboards to give them an idea of potential configurations (including the relative proportion that would include adjacent shapes). They then played a single practice game and completed a quiz that tested their understanding of the instructions, including when interference would occur. At the end of the experiment a single game was randomly selected to determine their bonus.

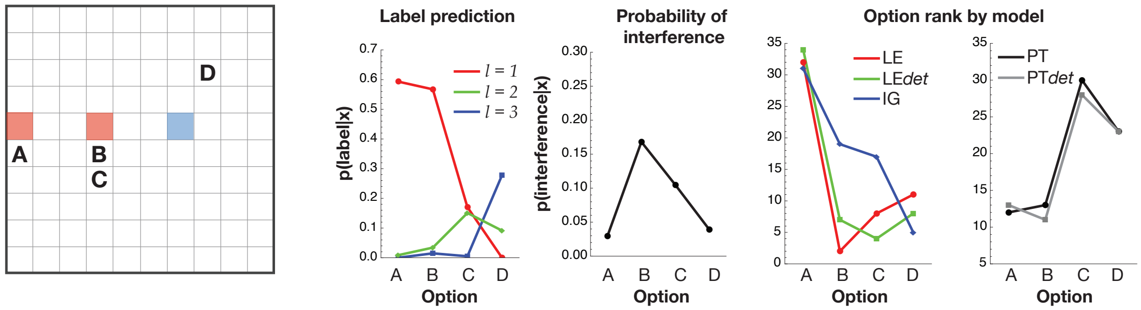

Figure 3: Depiction of the divergence between models after three observations have been made (two HITs on the red shape, one HIT on the blue shape). The locations in the grid, {A, B, C, D}, differ in their value according to each model, which is assessed here using the rank of each observation (with a rank of 1 indicating the observation that is preferred the most). For example, location A has a high probability of belonging to the red target (see "Label prediction" plot) and a low probability of interference (see "Probability of interference"). As a result, all of the uncertainty-based models (LE, LE_{det}, and IG) assign it a relatively low ranking, while the positive testing models assign it a high ranking. Notably, because of the low chance of interference, this location is ranked slightly higher by PT (which values receiving positive feedback during sampling) than PT_{det} (which is indifferent to the probability of interference). The remaining locations show how the model rankings change with the uncertainty in the outcome of the observation and the possibility of interference. Location B is highly likely to produce interference, and is ranked highest by LE, whereas location C has a lower probability of interference and is ranked highest by LE_{det}. Location D has a low chance of interference but still high uncertainty about its label, and is ranked highest by IG.

Results

Model comparison Participants could make up to 25 samples before their bonus reached \$0. For any games in which people sampled more than 25 times, we excluded those additional samples from analysis. We compared participants' choices to the predictions of five sampling models (PT, PT_{det}, LE, LE_{det}, and IG) by finding the best-fit temperature parameter β (see Eqn. 3). The number of participants that were best-fit overall by each model are shown in Figure 4A. Relatively few participants were best described by positive testing, with 6 participants best-fit by PT_{det} and only one participant best-fit by PT (these were combined for all following analyses). The greatest number of participants were best-fit by LE_{det} (16, or 37%), while 11 participants were best-fit by LE. The remaining 9 participants were best-fit by IG. Taken together, a majority of participants (63%) were best-fit by a model based on predicted label uncertainty (LE or LE_{det}).

Average bonus A one-way ANOVA on average bonus was performed with best-fit model as a between-subjects factor (see Figure 4B). There was a significant main effect of best-fit model ($F(39,3) = 12.4$, $p < .001$). Post-hoc independent samples t-tests were used to assess groups differences. PT participants achieved a significantly lower average bonus than all other groups (LE: $t(16) = -2.1$, $p = .05$; LE_{det}: $t(21) = -4.7$, $p < .001$; IG: $t(14) = -4.4$, $p < .001$). The LE group also performed significantly worse than the LE_{det} group ($t(25) = -4.1$, $p < .001$) and IG group ($t(18) = -3.6$, $p < .01$). There was not a significant difference between LE_{det} and IG participants ($t(23) = .5$, $p = .6$).

Reaction time Grouping participants by the best-fit model also revealed that participants differed in how long they took to query locations during the sampling phase (see Figure 4C). A one-way ANOVA on average median RT with best-fit model as a between-subjects factor revealed a significant main effect ($F(3,39) = 6.7$, $p < .001$). Post-hoc independent samples t-tests were then used to assess differences between groups. There was no difference in response time between PT and LE participants ($t(16) = -2.0$, $p = .06$), but PT participants made faster choices than LE_{det} ($t(21) = -3.3$, $p < .005$), and IG ($t(14) = -3.7$, $p < .005$) participants. LE participants also responded faster than LE_{det} ($t(25) = -2.5$, $p < .05$) and IG ($t(18) = -3.0$, $p < .01$). There was no difference between LE_{det} and IG groups ($t(23) = -.4$, $p = .7$). Thus, participants whose sampling behavior was best accounted for by LE_{det} or IG tended to both take longer to make sampling decisions and had higher performance in the test.

Sampling errors Finally, we measured the frequency of two types of sampling errors, which corresponded to choices that the ideal observer could predict the outcome with certainty given previous observations. *Positive* sampling errors indicated a query that was known to belong to a target, whereas *negative* errors indicated a query that was known to not belong to any target. A Kruskal-Wallis one-way analysis of variance revealed an effect of best-fit model on the proportion of positive errors ($F(3,39) = 17.7$, $p < .001$) (see Figure 4D, left). Wilcoxon rank sum tests were used to evaluate pairwise differences between groups. As expected, participants best-fit by the PT model made a greater proportion of confirmatory sampling errors than LE ($z = 2.8$, $p < .01$), LE_{det} ($z = 3.7$, $p < .001$), and IG ($z = 3.0$, $p < .005$) participants. There were no differences in the proportion of positive errors between the LE, LE_{det}, and IG groups (all $p > .05$). In contrast to positive errors, there was no overall difference between groups in the proportion of negative errors ($F(3,39) = 3.5$, $p = .3$). The low rate of negative errors in general suggests that participants took previous observations into account when making sampling decisions (e.g., a random sampling strategy would be expected to have a high rate of such errors).

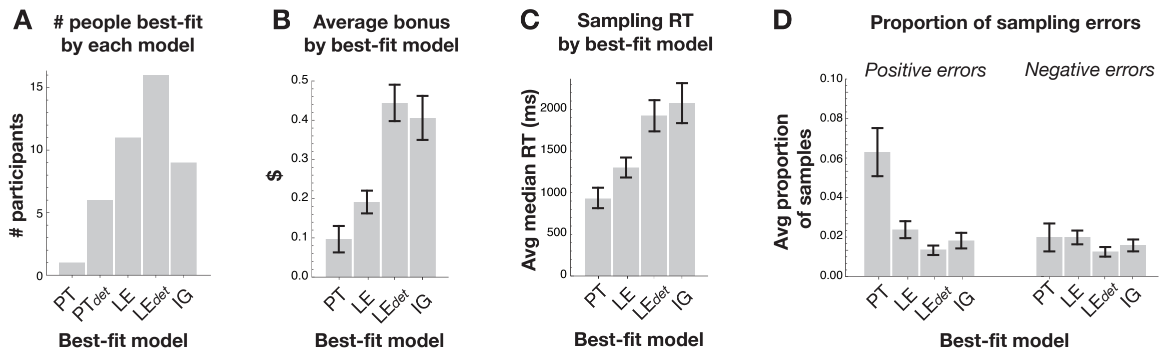

Figure 4: Results of the experiment.

Discussion

Normative theories like IG provide a useful framework for evaluating information search across different domains, but do not specify the process underlying sampling decisions. Like previous examples of positive testing being consistent with IG, under deterministic hypothesis spaces predictive uncertainty is equivalent to IG. The LE model represents a plausible process (assessing confidence in the prediction of an outcome) that often leads to optimal decisions, but which is less efficient under conditions of varying reliability. As a result of the noisy feedback process in our task, LE was not an adaptive sampling strategy, as it predicted a preference for locations where the probability of interference was high and less information about the true underlying label would be conveyed. Although participants were tested on their understanding of when interference would occur prior to beginning the experiment, it did not seem to influence sampling decisions in the majority of participants. Indeed, most people were best-described by a form of predictive uncertainty, with the largest group of participants best-fit by LE_{det}, which valued uncertainty about how to predict the true label of an observation while ignoring its reliability.

One surprising aspect of the results was that participants best-fit by LE_{det} and IG achieved the same average bonus across games, despite the fact that IG should lead to better performance. This lack of a difference was not attributable to either group committing more errors during the test phase, as the model's expected number of errors based on the information participants uncovered during sampling was in line with their performance. Although 2/3 of the games each participant played included adjacent shapes, one possibility is that interference still wasn't common enough to produce a significant advantage for those participants who took it into account when making sampling decisions.

Our finding that few participants were best-fit by IG refines the conclusion offered in our previous paper (Markant & Gureckis, 2012) which found that participants were better fit by IG than by a model that maximizes expected economic utility. This result raises the question of how people might learn to make better sampling decisions in this kind of task. Unlike perceptual decision making tasks in which the reliability of a stimulus is directly available (e.g., Orhan, Michel, & Jacobs, 2010), during information search the reliability must be estimated (e.g.,, in our task, by predicting the likelihood of shapes being adjacent). This estimation may improve with more experience, whereby people learn that some kinds of queries (e.g., those falling between two squares from different targets, as in Figure 2, trial 5) tend to produce less useful feedback. An interesting question for future work is how this kind of experience-based marker of information value might interact with simpler forms of predictive uncertainty to produce better information selection decisions (Gottlieb et al., 2013).

Acknowledgments. This work was supported by grant number BCS-1255538 from the National Science Foundation and the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract D10PC20023 to TMG.

References

Gottlieb, J., Oudeyer, P., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593.

Gureckis, T., & Markant, D. (2009). Active learning strategies in a spatial concept learning game. In N. Taatgen, H. van Rijn, L. Schomaker, & J. Nerbonne (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 3145–3150). Austin, TX: Cognitive Science Society.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*(2), 211–228.

Markant, D., & Gureckis, T. M. (2012). Does the utility of information influence sampling behavior? In N. Miyake, D. Peebles, & R. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 719–724). Austin, TX: Cognitive Science Society.

Najemnik, J., & Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391.

Navarro, D., & Perfors, A. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review*, *118*(1), 120.

Nelson, J. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *112*(4), 979–999.

Nelson, J., & Cottrell, G. (2007). A probabilistic model of eye movements in concept formation. *Neurocomputing*, *70*(13-15), 2256–2272.

Orhan, A., Michel, M., & Jacobs, R. (2010). Visual learning with reliable and unreliable features. *Journal of vision*, *10*(2).

Renninger, L., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision*, *7*(3).