Active learning as a means to distinguish among prominent decision strategies

Paula Parpart (paula.parpart.10@ucl.ac.uk), **Eric Schulz (eric.schulz.13@ucl.ac.uk)**,
Maarten Speekenbrink (m.speekenbrink@ucl.ac.uk) & **Bradley C. Love (b.love@ucl.ac.uk)**

Department of Experimental Psychology, University College London, London, WC1H 0AP

Abstract

A long-standing debate in decision making has been whether people rely on very little information for making choices, or weigh and add all available information. We propose a new method to determine whether a non-compensatory (Take-The-Best) or compensatory strategy (Logistic Regression) is more psychologically plausible: by looking at peoples active learning queries. This method goes beyond traditional model selection techniques as it reveals the information people choose to learn early on, which subsequently drives their decisions. We developed active learning algorithms for both Take-The-Best and Logistic Regression, and designed an active learning experiment to distinguish between these models. By letting both models and humans actively learn, we could compare their queries, and found that people follow a rank-based learning strategy in non-compensatory environments, but prefer more certainty-based queries in compensatory environments. We argue that active learning studies provide a promising new methodology to distinguish among decision models.

Keywords: Decision Making, Heuristics, Active Learning, Take-The-Best

Introduction

How do we decide between two alternatives? This question is as fundamental to research in judgment and decision making as its answer is controversial (Todd & Gigerenzer, 2000). Whereas some cognitive scientists believe that people only require a few pieces of information to come up with good decisions (Marewski et al., 2010), others have described humans as integrating all the evidence for both alternatives to make a decision (Arkes et al., 1986) . One of the core questions in this debate concerns the way in which people look up and integrate information.

Imagine you have to decide between two restaurants to go for lunch. Both restaurants differ on several binary cues (for example, one is in walkable distance, the other is not). One decision strategy you could apply is to weigh the cues by their importance and add them up; this is called a weighted additive linear model (Payne et al., 1993, WADD). For each restaurant, WADD would compute a weighted sum, and the restaurant with the larger sum is chosen. Alternatively, people might prefer a simpler strategy and base their decision only on one cue. This is what the Take-The-Best heuristic (Gigerenzer & Goldstein, 1996, TTB) does: it creates a ranking order of the cues according to their validities, and chooses the restaurant that is preferred by the highest ranked cue. If that cue does not discriminate between restaurants, then the second cue is considered, and so forth. WADD is a *compensatory* strategy, whereas TTB is a *non-compensatory* strategy. Compensatory strategies have the property that a cue can be compensated for by combinations of subsequent cues and regression is a typical examples thereof. In contrast, the non-compensatory Take-The-Best heuristic ignores most

cues to make decisions, as the most powerful cue C_k can outweigh any combination of the subsequent cues $C_{k+1}, \ldots, C_{k+n}$ (Gigerenzer & Goldstein, 1999). Both strategies are known to perform best in matching environments that have the same properties, i.e., WADD performs best in a compensatory environment and TTB in a non-compensatory environment (Martignon & Hoffrage, 1999).

Traditional model testing approaches

The dispute over the two model classes is about their psychological plausibility as models for human decision making. A repeated argument has been that non-compensatory strategies are simpler and require less computational capacity and are therefore more plausible (Todd & Gigerenzer, 2000). Yet, the most common method of pitting compensatory and non-compenastory strategies against each other have been statistical simulations, showing that one outperforms the other in artificial environments. For example, multiple studies show that the simpler TTB can outperform the compensatory linear regression (Czerlinski et al., 1999) in various datasets. In contrast, other studies show that there is no strong reason to prefer TTB over other cognitive models as it does not perform noticeably better (Chater et al., 2003; Schulz et al., 2014).

All of these studies argue about the predictive accuracy of cognitive models; nevertheless - just because one class of models can beat another with better predictions, it does not follow that this class is necessarily a better psychological representation of what people actually do. Although whole research paradigms are dedicated to solving the question about whether people rely on simpler non-compensatory heuristics or complex integrative mechanisms, different methodologies currently in use to answer this question are scarce and homogeneous. Another method has been probing whether participants look up additional information (Newell et al., 2003). This is only a limited approach to the problem at hand since it is only ever possible to check for $k + 1$ look-ups given C_k cues presented so far and it has been argued that people look up additional information but do not use it (Marewski & Mehlhorn, 2011).

We propose active learning as a novel method to solve the dilemma of discriminating among compensatory and non-compensatory strategies as psychologically plausible decision models. What most of the previous studies have in common is that they study peoples decision making in static, passive and highly controlled experiments. In order to answer the crucial question about what information people hold in memory and how they look up knowledge when making decisions,

we believe one has to look at an earlier stage in the process –at the stage of learning the relevant information in the first place. We argue that stronger evidence for peoples use of either TTB or a WADD strategy comes from the way people actively acquire information, i.e. cue weights and cue orders, in the respective environments. If a cognitive agent has evolutionarily developed to prefer a certain class of models as her means to learn a cognitive representation in a particular environment, then the way she sequentially selects information should (at least partially) reflect this representation. For example, if an agent has come to apply TTB, then -intuitively- she should try to find the most important cue first as this will decrease her uncertainty maximally, and so forth. Using this way of re-creating the structure of a cognitive mechanism, it becomes possible to set up active learning algorithms for many different cognitive models over time.

Active Learning: Do people learn with respect to cue weights or cue orders?

The main idea of psychological theories of active learning is to describe a learning agent as optimally designing experiments (Chaloner & Larntz, 1989). That is, given that one has to find the true hypothesis out of many potential explanations as fast as possible, an agent assigns prior probabilities to each hypothesis according to some objective criterion such as the available frequency data or according to the subjectively judged plausibility of each hypothesis. Each possible outcome of any possible experiment can thus be considered, and a "preposterior analysis" (Raiffa & Schlaifer, 1961) of the ways in which each possible experimental outcome could modify beliefs about the hypothesis, can be conducted. The proposals for optimal experimental design (OED) work in an information gain-driven fashion and maximize an informational utility, which is typically a measure of how much beliefs have changed or how large the uncertainty reduction is. There has been a great deal of interest in both normative and descriptive questions surrounding human information acquisition. In a probabilistic framework, many OED models have been used to model human behavior on cognitive tasks such as feature learning (Griffiths & Austerweil, 2009), reward-specific information search (Meder & Nelson, 2012), and to assess the trade-off between exploration and exploitation (Knox et al., 2011).

In this current study, we want to assess to what extent different OED models match participants' behavior in an active learning experiment, but with the goal of distinguishing among decision models. We specifically designed the perfect environments for both model classes, i.e., a fully non-compensatory environment for TTB, and a fully compensatory environment for logistic regression (WADD strategy), and several environments in between these extremes. As there are no active learning counterparts to the two decision models yet, we focus on qualitative predictions. Consequently, we developed two entropy-minimizing learning algorithms, one for a cue-ranking strategy and one for a cue-weighing strategy. Next, we compare the models' a priori search queries

to the queries made by participants in an active learning task with pairwise comparisons. By letting people freely choose among pairwise comparisons, we can investigate whether people pick information such that they learn about cue orders, or instead learn cue weights directly as proposed by the active WADD strategy.

Active learning algorithms

Both active learning algorithms essentially rely on a one-step ahead greedy entropy minimization of their posterior estimates. Greedy algorithms always choose as the next observation that which currently promises to reduce the uncertainty about the learning model maximally.

Take-The-Best

The active learning algorithm for the Take-The-Best heuristic is based on a entropy reduction method that considers a distribution over all possible cue orders. Therefore, we generated all possible cue orders given the features, including those where some cues are unimportant or negatively correlated with the outcome[1]. Afterwards, we put a uniform prior over the cue orders via pseudo counts. Pseudo counts here can be seen as urns, where the number of balls represents the current probability of a given order. The way we assess (Shannon's) entropy is by standardizing the balls per cue order via the sum of all balls. Using the resulting values as a probability density distribution, our algorithm predicts new cases by generating a TTB output for each of the cue orders and then estimating the overall mean, weighted by every cue order's current probability. Using this approach, we can easily estimate the probability for a new test item to be a win or a loss and calculate the expected reduction in entropy over all cues given a new training item. Entropy can be reduced via some cue orders generating correct predictions and thereby getting more counts on their pseudocounts. This is possible as our model updates all cue orders that made correct predictions at time point t by adding one more ball to their urns at time point $t + 1$. The addition of balls will –over time– put more and more probability mass on the true cue order and thereby gradually reduce uncertainty. The used distribution over all possible cue orders can also be seen as a distribution over the whole hypothesis space. The way our algorithm then works is by the attempt to drive down the uncertainty of this hypothesis space as quickly as possible, an approach that is close to optimal in a Bayesian active learning context (Golovin et al., 2010)[2].

Logistic Regression

Logistic regression is set up as a competing compensatory model to the non-compensatory TTB model. We use a Bayesian version of logistic regression based on a random walk Metropolis MCMC algorithm. Again, the way the algorithm works is by maximizing the expected information gain

---

[1]This results in a very complex hypothesis space that increases exponentially with the number of features.

[2]A technical description can be found in the appendix online.

for each trial. However, as a logistic regression does not learn cue orders, the expected information gain is approximated by the combined variance of all weights, the β-estimates. This means that a Bayesian logistic regression is fitted to all past observations and the current variance for all βs is calculated and then compared to the variances that could be expected from a newly fit model given a new input. This uncertainty sampling-based algorithm provides us with a compensatory active learning algorithm for the scenarios at hand. Instead of trying to drive down the hypothesis space as quickly as possible, this algorithm tries to learn about the weights as precisely as possible in order to make good compensatory judgments.

Degrees of Compensatoriness

We are interested in the performance of the two proposed active learning models in environments with different "compensatoriness". Note that a non-compensatory environment can be defined as a WADD environment in which the β weights are exponentially decreasing. In order to create different degrees of "compensatoriness", we make use of a mathematical trick that allows us to rely on a single parameter to smoothly vary from compensatory to non-compensatory environments through a "stick breaking process". The generation would be of a set of 4 weights $\beta_{k=1}^4$ through:

$$\beta'_k \sim \text{Beta}(1, \theta) \tag{1}$$

$$\text{Define } \{\beta'_k\}_{k=1}^4 \text{ as:} \tag{2}$$

$$\beta_k = \beta'_k \prod_{i=1}^{k-1}(1 - \beta'_i) \tag{3}$$

As the expectation of the Beta-distribution is defined as $\frac{\alpha}{\alpha+\theta}$, a perfect Take-The-Best environment corresponds to setting θ to 1 or greater as this would lead to a perfectly non-compensatory weight structure. Given the strict boarder of $\theta = 1$ that separates compensatory from non-compensatory strategies, we will use $\theta = [\,0, 0.5, 1, 2, \infty]$ for all the upcoming scenarios as this generates degrees of compensatoriness starting from uniform weights ($\theta = 0$) all the way to an environment where only one cue matters ($\theta = \infty$). Figure 1 shows the weighting structures that result from simulating different levels of compensatoriness for four cues by increasing θ.

Experiment

The experiment was designed to find out whether people are more likely to follow a rank-based or a weight based active learning algorithm. The outcome has strong implications for either decision mechanism as plausible decision strategies. We hypothesized that people are sensitive to the structure of the environment (the degree of compensatoriness) in their choice of learning algorithms. We assigned people randomly to one of the five above-mentioned compensatoriness conditions. We expected participants in the more non-compensatory environments to be better matched
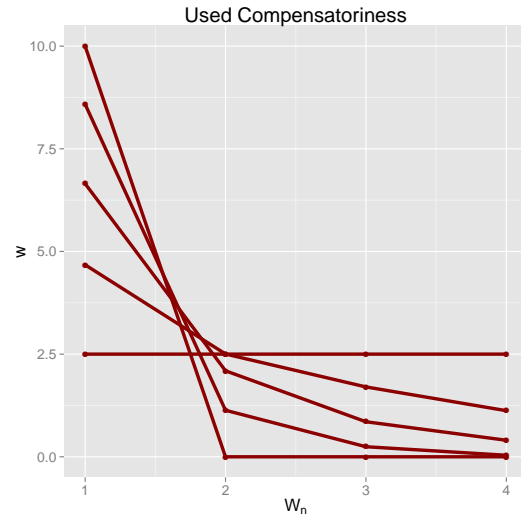

Figure 1: Compensatoriness for five different levels of θ. The x-axis represents four different cues and the y-axis displays the weight magnitudes.

by the active Take-The-Best algorithm, while logistic regression would better match their choices in the compensatory environments.

Participants

Two hundred and sixty-four participants ($N = 264$) participants ($M = 35.4$ years) were recruited via Amazon Mechanical Turk to take part in the "Alien Olympics" study. Participants were paid $0.50 for participation plus an additional bonus between $0 and $0.5 depending on their performance.

Materials and Stimuli

On each trial, participants had to choose a pair of Aliens to compete against each other, in order to learn about their strengths. The Aliens varied on four different features, which are displayed in Figure 2. The features were designed to be helpful in fights, e.g., wings enabled an Alien to fly which helps in attacking enemies, while camouflage is useful for defense. The features were explained to participants at the start and they were told that the different characteristics might not all be of equal importance for an Alien's strength in a fight. As there are four features, we generated all possible feature combinations which results in 16 different Alien types. On each trial, participants were presented with four random Aliens on the screen, and had to choose two of these to compete against each other. After selecting a pair, they received feedback about which Alien had won the competition. They were also told that sometimes a weaker Alien could win against a stronger competitor as in any sport, which reflects the probabilistic generation of the actual outcomes. The underlying weights of the four features that people could learn depended on the compensatoriness condition a participant was in. Importantly, we emphasized that people should

pick their Aliens wisely by selecting informative comparisons out of the presented Aliens, as the goal was to learn how the different features influenced an Aliens chances to win. Participants were informed that they would need this feature knowledge later in the experiment for an assessment task. The actual outcomes observed in feedback were generated by using the weights from the respective compensatoriness conditions (standardized to always add up to 10) and applying logistic regression in order to determine an Alien's strength, or likelihood of winning against another Alien.

Procedure

The experiment was divided into a learning phase and a test phase. The learning phase consisted of participants actively choosing Alien pairs to fight against each other on 30 trials. The test phase was designed to assess what people learned and was structured as follows: On each trial, participants were presented with only 2 different Aliens that were again randomly drawn from the Alien database. We told participants that these Aliens were the candidates for their Olympic Team, and it was their task to choose the Alien they considered to be stronger based on what they had learned about the characteristics. This assessment phase consisted of 10 binary choices. Participants were reminded that a bonus payment would depend on their performance in this test phase.
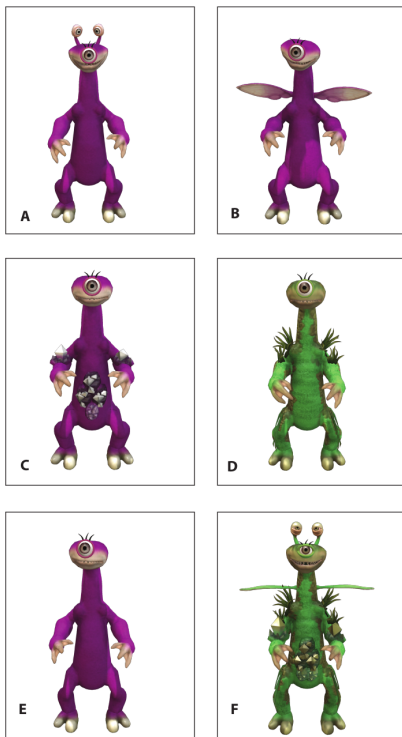

Figure 2: Aliens varied on 4 different features (A-D): Antennae, Wings, Diamonds, and Camouflage. E: Alien without features, F: Alien with all features.

Results

Participants' performance at identifying the stronger Aliens during the test stage was highly above chance; the average percentage of correct choices made was 74% ($t(263) = 27.44$, $p < 0.001$) with a range of (30%,97%). Performance varied as a function of the compensatoriness condition that participants were in. Figure 3 represents the average score as a function of compensatoriness: as the environmental structure gets more non-compensatory (i.e. more weight on just a few cues), the average performance drops. This intuitively makes sense as there is less information to be learned when a cue dominates all others, which makes draws more likely and informative comparisons less likely. However, peak performance was observed for an environment not completely compensatory ($\theta = 0$), but slightly compensatory.
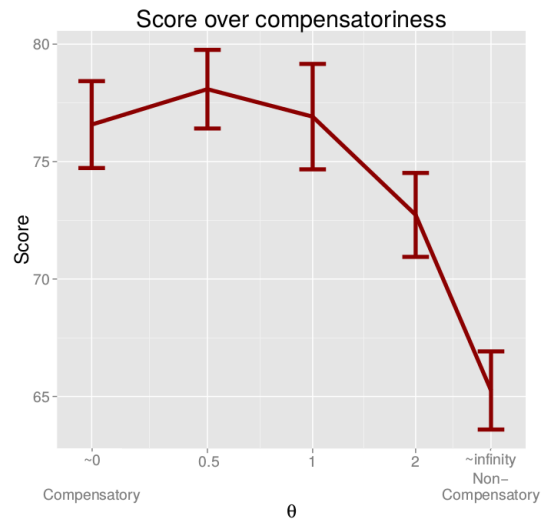

Figure 3: Average test performance across participants by compensatoriness conditions.

Next, we demonstrate the model fits of the regular Take-The-Best heuristic and Logistic Regression to participants' choices at test. These were generated by fitting both models to the set of Aliens participants had selected, and using the fitted models to predict choices in the test set. Figure 4 presents the model fits as a function of the compensatoriness. It can be seen that Take-The-Best was better at predicting people's behavior in the highly non-compensatory conditions, whereas in the more compensatory conditions no significant difference between the models could be found. This demonstrates again that purely behavioral model fits are limited in their ability to distinguish between common decision models, even in a prediction-based test. Therefore, we focus on the active learning results next.

Active Learning Results We categorized all possible pairwise comparisons people could choose on each trial into the 8 possible subtypes that can be seen on the x-axis of Figure 5.
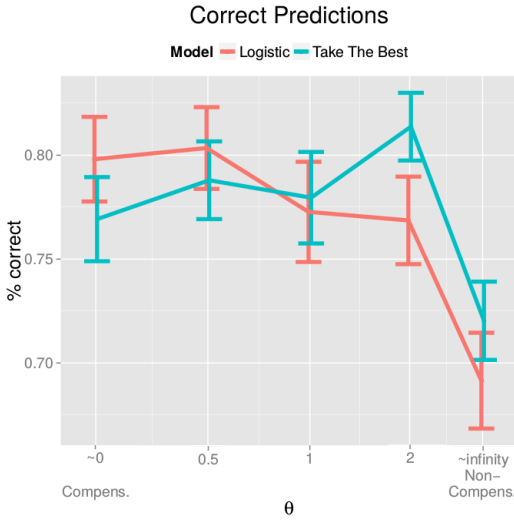
Figure 4: Behavioral model fits of the Take-the-Best Heuristic and Logistic Regression. The y-axis represents the percentage of correct predictions with respect to people's choices at test.

Figure 5: The 8 subtypes of active learning choices that participants made and the a priori model choices.

WDDD for example signifies a comparison of two Aliens with 3 similar features (D for draw), where one of the Aliens had one more feature than the other (W for Win). A WLDD comparison compares two Aliens with an equal direction of two features but differing in two others (W for Win, L for Loss), i.e., this might test whether the feature Wings are more important than Camouflage for the outcome. People rarely chose comparisons where it is unclear what feature was responsible for an outcome, e.g., a comparison of an Alien with 3 or 4 more features than its competitor (WWWD, WWWW). Instead, the most common comparisons were simpler and controlled, such as the WLDD, which test for the relative effect of one feature in comparison to another, i.e. searching for orders among features. Although participants were told in the instructions that all features are helpful in fights, the most frequent comparison was the WDDD query, assessing whether a feature improves the outcome. The fact that people preferred these simple queries could reflect a preference to perform confirmatory tests (Markant & Gureckis, 2012). Finally, we compared queries selected by the two learning algorithms against queries chosen by participants. We let both the TTB and Logistic Regression algorithms learn in the same compensatory and non-compensatory environment as the participants, by creating as many simulated participant profiles as there were participants in each compensatoriness condition. Then, we let the models learn over time. As a result, the frequencies of selections executed at each run by the algorithms were matched with those by humans and an overall correlation between the frequencies was calculated. Figure 6 shows in the lower panel how well the selection frequency predicted, and in the upper panel how well the selection frequencies by the models correlated with participants' frequencies.
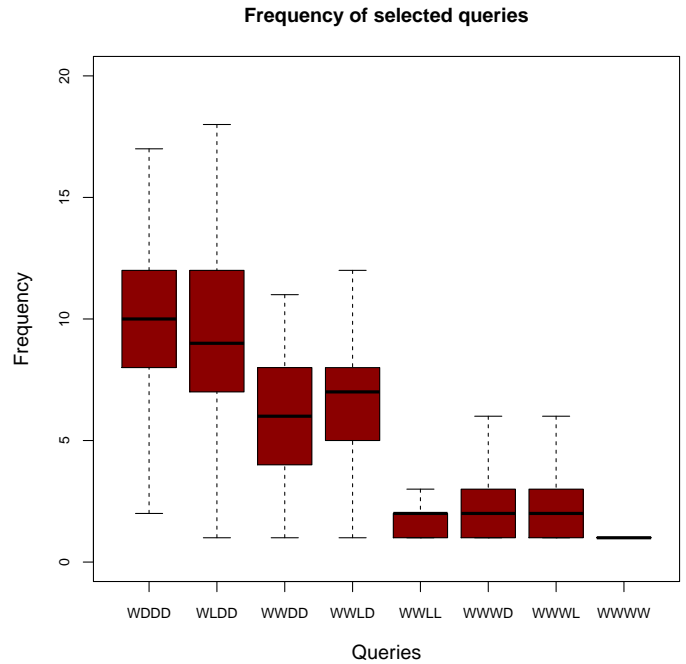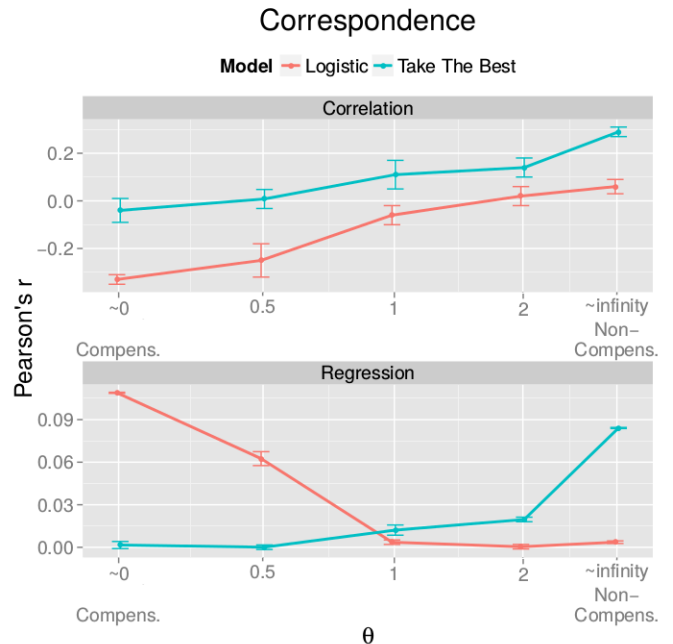

Figure 6: Frequency of queries by active learning algorithms matched to participants' queries. Lower panel: y-axis shows the variance explained. Upper panel: y-axis displays correlations.

These results demonstrate that participants in different con-

ditions differed in terms of which active learning algorithm best described their queries (lower panel, Figure 6). With an increase in non-compensatoriness, the Active Take-The-Best model explained more variance and Logistic Regression less, in line with our prediction. This is also reflected in the number of noncompensatory tests performed (e.g., WWLD), which increased in more non-compensatory conditions. The underlying correlations between the models' and peoples' selections (upper panel, Figure 6) reveal a similar picture for the TTB learning algorithm - with more non-compensatoriness, the TTB algorithm's queries better matched what queries people were choosing. However, as the environment becomes more and more compensatory, participants tend to select observations that do not reduce uncertainty maximally with respect to the weight-based logistic regression algorithm, and instead select comparisons that are more certain which could reflect more confirmatory hypothesis testing. This finding also remained the same when we used predictive uncertainty instead of information gain, showing that -as environments get more compensatory- participants tend to select queries where they already know the outcome better. Overall, the Take-The-Best active learning algorithm which minimizes uncertainty with respect to cue orders was a better description of how people learn in our experiment.

Discussion and Conclusion

Results demonstrated that–at least in more non-compensatory environments– people learn more like an order-based strategy, which lends support to the Take-The-Best heuristic as a plausible decision mechanism in these environments. This finding represents more processing-based evidence than the usual predictive accuracy findings or descriptive model fits. Furthermore, people do seem to be adaptive in their choice of active learning strategies to the structure of the environment: the more non-compensatory the environment was, the more choices people made in accordance with the TTB active algorithm. People were not well-described by the weight-based logistic regression algorithm, and for very compensatory environments they even seemed to pick the opposite choices to the uncertainty reducing algorithm, i.e., they chose more certain queries. This could in turn be due to the fact that –when all cues are equally important– people notice they cannot learn any cue rank orders, and so they might apply another strategy that is compensatory but not captured by logistic regression, for example a tallying strategy. This is a question we plan to address in a follow up study.

The current experiment demonstrates that active learning experiments can be used to distinguish among prominent decision strategies. Our results revealed a more informative picture than the traditional passive model fitting procedures. We believe that the application of active learning algorithms as a means to distinguish among decision algorithms is a novel and promising approach. Creating active versions of classic decision models and testing them in related experiments will

shed more light onto how people actually learn in different tasks. We hope to utilize this new approach for more process-based comparisons between different learning and decision making models in the near future.

References

Arkes, H. R., Dawes, R. M., & Christensen, C. (1986). Factors influencing the use of a decision rule in a probabilistic task. *Organizational Behavior and Human Decision Processes*, *37*(1), 93–110.

Chaloner, K., & Larntz, K. (1989). Optimal bayesian design applied to logistic regression experiments. *Journal of Statistical Planning and Inference*, *21*(2), 191–208.

Chater, N., Oaksford, M., Nakisa, R., & Redington, M. (2003). Fast, frugal, and rational: How rational norms explain behavior. *Organizational behavior and human decision processes*, *90*(1), 63–86.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics?

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, *103*(4), 650.

Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: take the best and its relatives. *Simple Heuristics that Make Us Smart. Oxford University Press, New York*, (pp. 75–95).

Golovin, D., Krause, A., & Ray, D. (2010). Near-optimal bayesian active learning with noisy observations. In *Advances in Neural Information Processing Systems*, (pp. 766–774).

Griffiths, T. L., & Austerweil, J. L. (2009). Analyzing human feature learning as nonparametric bayesian inference. In *Advances in neural information processing systems*, (pp. 97–104).

Knox, W. B., Otto, A. R., Stone, P., & Love, B. C. (2011). The nature of belief-directed exploratory choice in human decision-making. *Frontiers in psychology*, *2*.

Marewski, J. N., Gaissmaier, W., & Gigerenzer, G. (2010). Good judgments do not require complex cognition. *Cognitive processing*, *11*(2), 103–121.

Marewski, J. N., & Mehlhorn, K. (2011). Using the act-r architecture to specify 39 quantitative process models of decision making. *Judgment and Decision Making*, *6*(6), 439–519.

Markant, D., & Gureckis, T. M. (2012). One piece at a time: Learning complex rules through self-directed sampling. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society*.

Martignon, L., & Hoffrage, U. (1999). Why does one-reason decision making work. *Simple heuristics that make us smart*, (pp. 119–140).

Meder, B., & Nelson, J. D. (2012). Information search with situation-specific reward functions. *Judgment and Decision Making*, *7*(2), 119–148.

Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone takes-the-best. *Organizational Behavior and Human Decision Processes*, *91*(1), 82–96.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press.

Raiffa, H., & Schlaifer, R. (1961). Applied statistical decision theory. *Division of Research, Harvard Business School, Boston, MA*.

Schulz, E., Speekenbrink, M., & Shanks, D. R. (2014). Predict choice: A comparison of 21 mathematical models. Cognitive Science Society.

Todd, P. M., & Gigerenzer, G. (2000). Précis of simple heuristics that make us smart. *Behavioral and brain sciences*, *23*(05), 727–741.