Heuristics From Bounded Meta-Learned Inference

Marcel Binz

University of Marburg

Samuel J. Gershman

Harvard University

Eric Schulz

Max Planck Institute for Biological Cybernetics

Dominik Endres

University of Marburg

Author Note

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 290878970-GRK 2271, Project 2. Eric Schulz is supported by the Max Planck Society and the Jacobs Foundation. The authors would like to thank Björn Meder, Lion Schulz and Raphael Schween for comments on an earlier draft.

Correspondence concerning this article should be addressed to Marcel Binz, Department of Psychology, Theoretical Cognitive Science Group, University of Marburg, Gutenbergstraße 18, 35032 Marburg, Germany. E-mail: binz@staff.uni-marburg.de

Data and code available at: https://github.com/marcelbinz/HeuristicsFromBMLI

Preprint. This paper has not been peer reviewed or published yet.

Abstract

Numerous researchers have put forward heuristics as models of human decision making. However, where such heuristics come from is still a topic of ongoing debates. In this work we propose a novel computational model that advances our understanding of heuristic decision making by explaining how different heuristics are discovered and how they are selected. This model, called bounded meta-learned inference, is based on the idea that people make environment-specific inferences about which strategies to use, while being efficient in terms of how they use computational resources. We show that our approach discovers two previously suggested types of heuristics – one reason decision making and equal weighting – in specific environments. Furthermore, the model provides clear and precise predictions about when each heuristic should be applied: knowing the correct ranking of attributes leads to one reason decision making, knowing the directions of the attributes leads to equal weighting, and not knowing about either leads to strategies that use weighted combinations of multiple attributes. This allows us to gain new insights on mixed results of prior empirical work on heuristic decision making. In three empirical paired comparison studies with continuous features, we verify predictions of our theory, and show that it captures several characteristics of human decision making not explained by alternative theories.

Keywords: meta-learning, resource rationality, heuristics, strategy selection, strategy discovery

Heuristics From Bounded Meta-Learned Inference

Imagine having to decide which of two movies you are going to watch tonight: Movie A vs. Movie B. Movie A has a higher average rating on a website that you trust, while Movie B is directed by a known director and has previously won an Oscar for the best picture. From past experiences, you know that rating is the best indicator for a good movie. Whether the movie won an Oscar and who directed it is less important for how much you normally enjoy watching a movie. How do people make decisions like this?

The question of how people decide between two options is as fundamental as its answer is contentious. Indeed, even though we make countless such decisions every day, the underlying principles of these decisions are still debated in psychology (Todd & Gigerenzer, 2000), behavioral economics (Samuels et al., 2012), and neuroscience (Camerer et al., 2005).

Traditionally, researchers have approached this problem by looking at how rational agents decide. From this ideal observer perspective (Geisler, 1989) it is assumed that people weight different attributes of each option appropriately to combine information from all available sources. Psychologists were however quick to point out that rational decision making can be too burdensome (Simon, 1990b; Tversky & Kahneman, 1974). Instead, they suggested that human decision making may be based on a variety of heuristics, which are simple strategies that ignore part of the relevant information (Gigerenzer & Todd, 1999; Shah & Oppenheimer, 2008; Tversky & Kahneman, 1974).

Two common classes of heuristics are *one reason decision making* (Gigerenzer & Goldstein, 1999) and *equal weighting* (Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975). One reason decision making heuristics are based on the idea that good reasoning often requires just a single piece of information (Marewski et al., 2010). Applying such a strategy to the initial example, you would only need to inspect the most important attribute: the movie rating. Based on this attribute, you decide to watch Movie A and ignore all other information about both movies. Equal weighting heuristics on the other hand completely abstain from differentiating between the attributes and instead tally all of them together to decide which option to choose. In our example, Movie B has two attributes in its favor, while Movie A only has one. Hence, you would decide to watch Movie B if your decision was based on an equal weighting heuristic.

Even though they are computationally simplistic strategies, heuristics can be surprisingly competitive in many real-world benchmarks (Czerlinski et al., 1999; Lichtenberg & Şimşek, 2017). This observation led different researchers to consider heuristics as *ecologically rational* strategies (Gigerenzer & Gaissmaier, 2011; Gigerenzer & Todd, 1999; Payne et al., 1993), implying that heuristics are strategies that are particularly well-suited for our complex and dynamic world. The ecological rationality of heuristics also makes it appealing to view them as models of human decision making. Empirical studies attempting to show that people actually apply heuristics have however produced mixed evidence (Ayal & Hochman, 2009; Bröder, 2000; Bröder & Gaissmaier, 2007; Glöckner & Betsch, 2008; Hilbig, 2010, see also our later discussion on empirical results).

In this work we suggest *bounded meta-learned inference* (BMLI) as a novel computational theory for explaining how people make decisions. BMLI discovers decision making strategies through a resource rational inference algorithm (Gershman et al., 2015; Lieder & Griffiths, 2019; Simon, 1990a) that has been adapted to an environment over time via meta-learning (Bengio et al., 1991; Schmidhuber et al., 1996; Thrun & Pratt, 1998). Like ideal observer models, BMLI attempts to infer optimal decision making strategies, but does so while taking computational resources into account. Like heuristics, strategies inferred through BMLI are tailored to a specific environment. However, unlike heuristics, the inductive biases of such strategies have been meta-learned through previous interactions with the environment instead of building them in by design.

Through a series of model simulations we demonstrate that BMLI discovers several previously suggested heuristics. Specifically, our results reveal three important classes of environments that lead to three different strategies. First, if the model knows the correct ranking of attributes but not their weights, then it learns a strategy that makes decisions based only on the attribute with the highest ranking, a form of one reason decision making. Secondly, if the model knows that the direction of correlation between attributes and outcome is positive, then it learns a strategy that makes decisions based on equal weighting. Finally, if the model does not know either the ranking nor the direction of attributes, it learns to use individual weights for each attribute. This analysis provides new insights on the mixed results of prior empirical work on heuristics, because it makes precise predictions about if and when a specific heuristic should be used. We verify these predictions in three empirical paired comparison studies and show that the vast majority of participants apply heuristics whenever they are optimal strategies for the current environment after taking limited computational resources into account.

In summary, our work makes the following three main contributions:

1. We show that heuristics can emerge through BMLI, thereby providing a normative justification for previously suggested heuristics.

2. We clearly map out which features of an environment lead to which (heuristic) decision making strategy, where knowing the correct ranking of attributes leads to one reason decision making, knowing the directions of the attributes leads to equal weighting, and not knowing about either leads to strategies that use weighted combinations of multiple attributes.

3. We test these predictions empirically in three experiments and find strong evidence for our theory's predictions, thereby reconciling several past contradictory results.

The remainder of the paper is organized as follows: we first summarize the relevant literature on heuristic decision making and introduce its general terminology. Afterwards, we present formal models corresponding to different hypotheses considered in our work. By running simulations on different environments, we generate several predictions of our theory, which we empirically test in three new decision making experiments. Finally, we discuss our results and connect our theory to related ideas.

Past Research on Heuristic Decision Making

There has been an extensive amount of past research on heuristic decision making. In this section, we describe common heuristics and review prior studies in the paired comparison setting with a focus on the empirical evidence for heuristic decision making.

Heuristics Toolbox

Although a mathematically precise definition of what constitutes a heuristic is still a topic of ongoing debates (Chater et al., 2003; Van Rooij et al., 2012), here we adopt the following definition put forward by Gigerenzer and Gaissmaier (2011): "A heuristic is a strategy that ignores part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods."

The collection of different heuristics is often thought of as an adaptive toolbox from which appropriate decision making strategies can be selected as required (Gigerenzer & Selten, 2002). We are primarily interested in heuristics that can be applied to paired comparison tasks (e.g., Martignon & Hoffrage, 2002) like the aforementioned movie example. In such tasks, a decision making agent is asked to judge which of two options is superior on an unobserved criterion. To aid the decision making process, the agent observes multiple attributes of both options, also known as cues or features in the decision making literature. Most heuristics developed for the paired comparison setting make use of binary features that indicate whether an attribute is present or not.[1]

Many heuristics are built around the concept of feature validity (Todd & Dieckmann, 2005). The validity of a binary feature is the rate at which it allows the agent to make correct predictions given that the feature is present in one option but not the other (Lee & Cummins, 2004). For example, the validity of being directed by a known

[1] Note that non-binary features, like average movie ratings, can always be dichotomized at a loss of information. In past studies, this has been frequently done by setting values which were less than the median to 0 and otherwise to 1.

director for predicting whether you like a movie could be 0.8, indicating that you would enjoy a movie that is directed by someone you know over someone you do not in eighty percent of the cases. In general, decision making strategies for paired comparison tasks can be categorized into two classes: compensatory and non-compensatory strategies. A strategy is compensatory whenever it integrates information from multiple features, whereas it is non-compensatory when a feature cannot be outweighed by any combination of less important features (Rieskamp & Hoffrage, 1999).

The weighted additive (WADD) strategy (Gigerenzer & Goldstein, 1996) is an example for compensatory decision making. WADD weights features by their validities, and decides for the option with the larger sum of weighted features. Although WADD combines information from multiple sources, it is – according to our definition – a heuristic, because it ignores potential interactions between features. In our movie example, this would correspond to weighting and adding all features together without paying attention to how they might interact (e.g., a movie data base could potentially always dislike Oscar-winning movies for being too mainstream; WADD would ignore this interaction).

Most heuristics are however much simpler than WADD. Equal weighting heuristics, for example, are compensatory, yet simple, decision making strategies. They do not distinguish between how features are weighted and instead use an identical weighting for all features. The process itself is realized by tallying features of both options together and selecting the one with the larger sum (Dawes & Corrigan, 1974; Einhorn & Hogarth, 1975).

The prime example for a non-compensatory strategy is the take-the-best (TTB) heuristic (Gigerenzer & Goldstein, 1996). TTB belongs to the family of one reason decision making heuristics. It assumes a ranking of features based on their validities and inspects features in decreasing order until a feature that discriminates between both options is reached. The final decision is based on the validity of that feature alone, ignoring all other information. Should a ranking of features not be a priori accessible, then it can either be estimated from observations or a random ranking can be used. A TTB strategy using a random ranking of features is referred to as the Minimalist heuristic (Gigerenzer & Goldstein, 1996).

Empirical Studies

The observation that heuristics are computationally efficient and ecologically rational strategies is often used to justify them as models of human decision making (Todd & Gigerenzer, 2007). However, to truly establish that people actually use heuristics, proving good performance in simulation is not sufficient; it also requires empirical evidence. Many studies have attempted to find such evidence, yet no consensus for or against heuristics has been reached. Here, we provide an overview of these studies and attempt to connect their findings. We consider studies in which information about features was freely accessible and those that included a cost for obtaining information. Tables 1 and 2 summarize characteristics of studies with and without costs to acquire information, respectively.

Problems where it is costly to access feature values naturally favor strategies that only require a few pieces of information. Because of that, studies in this context concentrated on one reason decision making heuristics such as TTB. For our review, we consider studies in the Mouselab paradigm (Payne et al., 1988), studies with monetary costs (Newell et al., 2003) and memory-based retrieval studies (Bröder & Schiffer, 2003).

The Mouselab paradigm is a process tracing approach to decision making, which requires participants to click or hover over a specific feature to reveal its value. This paradigm allows researchers to identify which information is considered by the participant. In studies making use of the paradigm, Rieskamp and Otto (2006) showed that people's selection of strategies depended on the environment they interacted with. Participants in their study had initial preferences for compensatory strategies, but then slowly adopted TTB in a non-compensatory environment and WADD in a compensatory one. However, other studies with comparable conditions arrived at different conclusions. For example,

Paper	Learning	Cost	Tasks	Trials/Task	Options	Features	Discretized	Ranking	Direction	Evidence
Rieskamp and Otto (2006, Study 1)	✗	↖	1	168	2	6	✓	✓	+	✓
Rieskamp and Otto (2006, Study 2)	✓	↖	1	182	2	6	✓	✗	+	✓
Glöckner and Betsch (2008, Study 2)	✗	↖	1	138	3	3	✓	✓	+	✓
Scheibehenne et al. (2013)	✗	↖	1	48	2	6	✓	✓	+	✗
Van Ravenzwaaij et al. (2014, Study 1)	✗	↖	1	100	2	9	✓	✓	+	✗
Van Ravenzwaaij et al. (2014, Study 2)	✗	↖	1	100	2	9	✓	✓	+	✗
Bröder (2000, Study 3)	✗	$	1	120	2	4	✓	✓	+	✓
Rieskamp and Otto (2006, Study 3)	✗	$	1	168	2	6	✓	✓	+	✓
Dieckmann and Rieskamp (2007)	✗	$	1	96	2	6	✓	✓	+	✓
Newell et al. (2003, Study 1)	✓	$	1	60	2	6	✓	✓	+	✗
Newell et al. (2003, Study 2)	✓	$	1	60	2	2	✓	✓	+	✗
Newell and Lee (2011, Study 2)	✓	$	1	80	2	6	✓	✗	+	✗
Bröder and Schiffer (2003)	✗	💾	1	52	10	4	✓	✓	?	✓
Bröder and Schiffer (2006)	✗	💾	1	52	10	4	✓	✓	?	✓
Bröder and Gaissmaier (2007)	✗	💾	1	52	10	4	✓	✓	?	✓

Table 1

Empirical studies that involved costs to acquire information about features. The learning column indicates whether validities/weights were provided (✗) or had to be learned (✓). The cost column describes how costs to acquire features were implemented, with ↖ referring to a Mouselab paradigm, $ indicating that monetary fees are required to reveal a feature and the floppy disk representing a memory-based retrieval setting. The direction column shows the direction of features, with + for positive directions and ? for unknown directions. The evidence column indicates whether the study found evidence for heuristics (✓) or not (✗).

Paper	Learning	Cost	Tasks	Trials/Task	Options	Features	Discretized	Ranking	Direction	Evidence
Bergert and Nosofsky (2007, Study 1)	✓	✗	1	160	2	6	✓	✗	+	✓
Bergert and Nosofsky (2007, Study 2)	✓	✗	1	160	2	6	✓	✗	+	✓
Bröder (2000, Study 1)	✓	✗	1	30	2	5	✓	✗	+	✗
Bröder (2000, Study 2)	✓	✗	1	120	2	5	✓	✗	+	✗
Lee and Cummins (2004)	✓	✗	1	5	2	6	✓	✗	+	✗
Glöckner and Betsch (2008, Study 1)	✗	✗	1	138	3	3	✓	✓	+	✗
Newell and Lee (2011, Study 1)	✓	✗	1	40	2	6	✓	✗	+	✗
Parpart et al. (2017)	✓	✗	1	10	2	4	✓	✗	+	✗
This work (Study 1)	✓	✗	30	10	2	4	✗	✓	?	✓
This work (Study 2)	✓	✗	30	10	2	4	✗	✗	+	✓
This work (Study 3)	✓	✗	30	10	2	4	✗	✗	?	✗

Table 2

Empirical studies that involved no costs to acquire information about features. The learning column indicates whether validities/weights were provided (✗) or had to be learned (✓). The direction column shows the direction of features, with + for positive directions and ? for unknown directions. The evidence column indicates whether the study found evidence for heuristics (✓) or not (✗).

Scheibehenne et al. (2013) demonstrated that people were better described through a mixture of TTB and WADD even in non-compensatory environments, indicating a general preference for compensatory strategies. Van Ravenzwaaij et al. (2014) showed that hierarchical models accounting for both search order and termination provided a better explanation for participants' choices than TTB and WADD.

Requiring a monetary cost to reveal features is another process tracing approach. Like the Mouselab paradigm, it facilitates strategies that rely on less information. In several experiments with monetary costs, Bröder (2000) produced evidence in favor of one reason decision making heuristics. In his experiments, more participants were classified as TTB users in a high cost condition compared to a low cost condition. Similarly, Dieckmann and Rieskamp (2007) observed that TTB predicted more decisions in environments with monetary costs. However, Newell et al. (2003) demonstrated that even with large monetary costs and other conditions favoring one reason decision making heuristics, not many participants acted according to TTB.

Requiring participants to recall features from memory is yet another method to constrain the amount of information they use. In multiple experiments with memory-based retrieval, Bröder and colleagues demonstrated that participants became more consistent with TTB when features had to be retrieved from memory (Bröder & Gaissmaier, 2007; Bröder & Schiffer, 2003, 2006). Bröder and Schiffer (2003), for example, classified 72% of participants as TTB users when they were under high working memory load, but only 56% when they were not.

In general, studies with increased costs for utilizing information indicate that human decision making becomes more consistent with one reason decision making heuristics. Nonetheless, even under supposedly favorable conditions, prior research did not reach a clear consensus on whether people use one reason decision making heuristics or if they rely on more complex strategies instead.

Glöckner and Betsch (2008) argued that process tracing studies are likely to underestimate the cognitive capacity of participants, as they hinder the activation of automatic decision making processes. They verified this claim by demonstrating that participants were generally able to combine information from multiple features extremely quickly when the acquisition of information was not constrained. Further studies with freely accessible information provided similar results (Bröder, 2000; Lee & Cummins, 2004; Parpart et al., 2018), always concluding that few participants made decisions consistent with TTB and that their choices were in general better described through compensatory strategies such as logistic regression. Newell and Lee (2011) highlighted large inter-individual differences and presented a sequential sampling model providing better fits than TTB, WADD and a strategy selection model across all participants. Bergert and Nosofsky (2007) were among the few who provided support for heuristics in human decision making even when information is free. They showed that people exhibit non-compensatory patterns of decision making, assigning over half of the total weight to a single feature and provided additional evidence for frugal strategies in form of reaction times.

To summarize, many past studies attempted to produce evidence for one reason decision making, thereby focusing less on other heuristics such as equal weighting. Many of them concluded that heuristic strategies were indeed more apparent when it was costly to access information. Evidence for heuristics in human decision making in the unimpeded setting is however rare. Looking at Tables 1 and 2, we observe that the majority of prior empirical studies evaluated their hypothesis in environments that either explicitly or implicitly assumed positive directions of features. While it is always possible to code features such that they have positive directions (e.g., changing the feature “won an Oscar” to “did not win an Oscar” if winning an Oscar has a negative correlation with the outcome), doing so can influence the strategies people apply. To foreshadow our results, we demonstrate that a restriction to environments with known positive attribute directions causes equal weighting heuristics to become optimal under limited computational resources. Therefore, at least some of the mixed results of prior studies can be explained by the use of environments that favor strategies not considered in their analyses.

Computational Models

At the heart of ecologically rational heuristics is a powerful idea: there are much simpler ways of interacting effectively with many natural problems than the use of complex strategies. Moreover, computational constraints render it necessary that people should make use of these shortcuts extensively. In our summary of previous empirical studies, we have seen that this is indeed sometimes observed, but certainly not always.

In this section we propose a new theory based on the idea that people make environment-specific inferences about which strategies to use, while at the same time making efficient use of their computational resources. To formalize and test this conjecture, we also introduce several other computational models of decision making in paired comparison tasks. First, we will outline the assumptions about the structure of the problem to be solved and define a corresponding ideal observer model. Then, we will introduce probabilistic variants of two popular heuristics. Both heuristics are considerable simplifications with respect to how they use information compared to the ideal observer model. Finally, we will describe how we obtain resource rational inference algorithms that are adapted to a particular environment. Having access to such a model will subsequently allow us to predict when and if people should rely on heuristic decision making strategies.

Ideal Observer

Ideal observer models (IO) are designed to provide a theoretical upper bound on performance in a specific task. In the following we construct an ideal observer model for paired comparison tasks.

In paired comparison tasks, an agent has to decide which of two options with feature vectors $\mathbf{x}_{A,B} \in \mathbb{R}^D$ has the higher value on an unobserved criterion $y_{A,B}$. In our movie example, the feature vector contains information about whether the movie has won an

Oscar, its average rating on a reviewing website and so on, while the unobserved criterion corresponds to your personal rating of the movie (i.e., how much you would like the movie).

We assume that the underlying relationship between features and the criterion is linear:

$$\begin{aligned} y_A &= \mathbf{w}^T \mathbf{x}_A + \epsilon_A \\ y_B &= \mathbf{w}^T \mathbf{x}_B + \epsilon_B \end{aligned} \tag{1}$$

with feature weights $\mathbf{w} \in \mathbb{R}^D$ and independent, additive noise $\epsilon_{A,B} \sim \mathcal{N}(0, \sigma^2)$. Under this assumption we can express the probability, that option A has a higher criterion value than option B as:

$$\begin{aligned} p(Y_A > Y_B | \mathbf{x}_A, \mathbf{x}_B, \mathbf{w}, m = \mathrm{IO}) &= p(C = 1 | \mathbf{x}, \mathbf{w}, m = \mathrm{IO}) \\ &= \mathbf{\Phi}\left(\frac{\mathbf{w}^T \mathbf{x}}{\sqrt{2}\sigma}\right) \end{aligned} \tag{2}$$

where $\mathbf{\Phi}$ is the cumulative distribution function of a standard normal distribution. For ease of notation we have denoted the difference between feature vectors as $\mathbf{x} = \mathbf{x}_A - \mathbf{x}_B$ and used a binary variable C that takes the value of 1 if $y_A > y_B$ and 0 otherwise.

Equation 2 makes it clear that an ideal observer should represent the probability that one option is better than the other using a weighted sum of differences between features of the options. Hence, the ideal observer model is a compensatory decision making strategy.

Heuristics

The two heuristics we consider in our analysis belong to the categories of one reason decision making and equal weighting. In contrast to traditional heuristics, like TTB, they are probabilistic decision making strategies for tasks with continuous features. Both are obtained through modification of the ideal observer model, such that either less information is required to make a decision or that information is combined in a simpler way.

One Reason Decision Making

In our implementation of one reason decision making we modify Equation 2 and replace it with a model that only takes a single feature $\mathbf{x}_*$ into account:

$$p(C = 1|\mathbf{x}, w, m = \text{SC}) = \mathbf{\Phi}\left(\frac{w \cdot \mathbf{x}_*}{\sqrt{2}\sigma}\right) \tag{3}$$

We refer to the resulting strategy as single cue heuristic (SC). If a ranking of features is available, decisions are based on the most predictive feature, otherwise we infer the most predictive feature from observations. In contrast to TTB, the single cue heuristic does not involve sequential search over features. However, we assume that features take continuous values, and hence search is not required as a feature nearly always discriminates between options (Luan et al., 2014).

Equal Weighting

In our probabilistic version of equal weighting, we replace Equation 2 with a model that has a single, tied weight for all features:

$$p(C = 1|\mathbf{x}, w, m = \text{EW}) = \mathbf{\Phi}\left(\frac{w \cdot \sum_{i=1}^{D} \mathbf{x}_i}{\sqrt{2}\sigma}\right) \tag{4}$$

If $w > 0$ this equal weighting heuristic probabilistically selects the option with the larger sum of features. For $w < 0$ it becomes more likely to select the option with the smaller sum. Using a negative weight is appropriate if most features have negative correlations with the criterion.

Parameter Estimation

We assume that the underlying weights $\mathbf{w}$ are not provided to the decision making agent. Instead, parameters of all models have to be estimated based on past observations. The ideal observer model contains as many free parameters as there are observed features.

Both single cue and equal weighting heuristics have only a single free parameter regardless of how many features are observed.

Let $(\mathbf{x}_1, c_1, \ldots, \mathbf{x}_T, c_T)$ represent all feature-target pairs from a single task. We view parameters as latent variables and estimate them by applying Bayesian inference sequentially. Exact inference is not possible under the above assumptions and thus we resort to a variational approximation (Jordan et al., 1999). The true posterior is approximated with a normal distribution $q(\mathbf{w}; \boldsymbol{\lambda}_t) = \mathcal{N}(\mathbf{w}; \boldsymbol{\mu}_t, \boldsymbol{\Psi}_t)$ and its parameters $\boldsymbol{\lambda}_t = (\boldsymbol{\mu}_t, \boldsymbol{\Psi}_t)$ are optimized through gradient ascent on the evidence lower bound:

$$\mathcal{L}(\boldsymbol{\lambda}_t) = \mathbb{E}_{\mathbf{w} \sim q(\mathbf{w}; \boldsymbol{\lambda}_t)} \left[\log p(C_t = c_t | \mathbf{x}_t, \mathbf{w}) \right] - \text{KL} \left[q(\mathbf{w}; \boldsymbol{\lambda}_t) || q(\mathbf{w}; \boldsymbol{\lambda}_{t-1}) \right] \quad (5)$$

where $q(\mathbf{w}; \boldsymbol{\lambda}_0)$ corresponds to an initial prior distribution and t denotes the number of feature-target pairs observed before the estimation. This kind of approximation is equivalent to exact inference when the true posterior is within the considered variational family. We provide further details on how Equation 5 is optimized in Appendix A.

In order to make predictions, we average over all plausible parameter values given by the variational distribution. The resulting predictive distribution can be expressed in closed form:

$$p(C_{t+1} = 1 | \mathbf{x}_{t+1}, \boldsymbol{\lambda}_t) = \int p(C_{t+1} = 1 | \mathbf{x}_{t+1}, \mathbf{w}) q(\mathbf{w}; \boldsymbol{\lambda}_t) d\mathbf{w} = \boldsymbol{\Phi} \left(\frac{\boldsymbol{\mu}_t^T \mathbf{x}_{t+1}}{\sqrt{2\sigma^2 + \mathbf{x}_{t+1}^T \boldsymbol{\Psi}_t \mathbf{x}_{t+1}}} \right) \quad (6)$$

The single cue heuristic additionally requires the identity of the most predictive feature. Whenever this is not available, we fit separate models for each feature and select the one with the highest accumulated evidence lower bound (Equation 5) over all past time steps.

Bounded Meta-Learned Inference

Next, we introduce bounded meta-learned inference (BMLI) as a novel theory for human decision making. BMLI is a resource rational inference algorithm that has been adapted to an environment over time via meta-learning. Both meta-learning and resource rationality are equally important in our approach, as we will later show by contrasting BMLI with a modification without resource limitations (MLI).

Meta-learning (Bengio et al., 1991; Schmidhuber et al., 1996; Thrun & Pratt, 1998), also known as learning to learn, is a machine learning approach to devise learning systems that can rapidly adapt to new problems. Specifically, we meta-learn an inference algorithm for paired comparison tasks. The idea is simple: instead of using Bayesian inference to infer posterior distributions over weights, we train a recurrent neural network to estimate distributions that are optimal for making future predictions.

The recurrent network takes a sequence of feature-target pairs $(\mathbf{x}_1, c_1, \ldots, \mathbf{x}_T, c_T)$ from a paired comparison task as inputs and computes a distribution over weights for each step – i.e., $\boldsymbol{\mu}_t$ and $\boldsymbol{\Psi}_t$. The estimated weights are combined with the next feature vector $\mathbf{x}_{t+1}$ to obtain the predictive distribution (using Equation 6). This process is illustrated graphically in Figure 1.

Through repeated encounters with an environment, the model adapts to properties of that specific environment.[2] Once training is complete, the recurrent network acts as free-standing inference algorithm without requiring any further optimization. Instead, adaptation to new tasks is realized through updates in its hidden activations, i.e., the network has learned to map past data to a distribution over parameters and thus acts as an inference algorithm itself. Here, we are interested in the properties of this emerging inference algorithms.

In addition to achieving good performance, we would like to obtain inference algorithms with low complexity. Let $\boldsymbol{\Theta}$ denote the neural network parameters, which we refer to as meta-parameters in order to distinguish them from the regression weights $\mathbf{w}$ of Equation 2. Meta-parameters are adapted through gradient descent on a loss function, which controls the trade-off between the accuracy of the network and the statistical

[2] For our purpose, an environment represents the distribution over tasks that can be encountered.

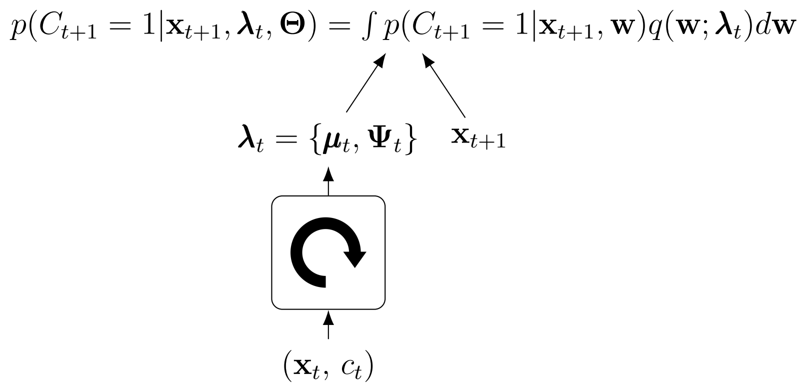

Figure 1

Graphical depiction of BMLI. The recurrent neural network sequentially processes examples from a given task. Through its recurrent activations it combines information from all previous feature-target pairs to compute a distribution over weights, which is then combined with the next input to obtain the predictive distribution.

complexity of meta-parameters, as measured by the entropy of an encoding distribution $q(\boldsymbol{\Theta}; \boldsymbol{\Lambda})$ relative to a prior (Achille & Soatto, 2018, 2019):

$$\mathcal{L}(\boldsymbol{\Lambda}) = \underbrace{-\mathbb{E}_{q(\boldsymbol{\Theta};\boldsymbol{\Lambda})}\left[\sum_{t=1}^{T} \log p(C_t = c_t|\mathbf{x}_t, \boldsymbol{\lambda}_{t-1}, \boldsymbol{\Theta})\right]}_{\text{error cost}} + \beta \underbrace{\mathrm{KL}\left[q(\boldsymbol{\Theta}; \boldsymbol{\Lambda})||p(\boldsymbol{\Theta})\right]}_{\text{complexity cost}} \tag{7}$$

where β controls the trade-off between performance and simplicity of the resulting inference algorithm, i.e. it determines to which degree computational complexity is taken into account.

Training recurrent neural networks with Equation 7 until convergence leads to BMLI. The objective in Equation 7 is closely related to performing variational inference over meta-parameters. The connection between variational inference and the Minimum Description Length principle (Grünwald & Grunwald, 2007; Hinton & Van Camp, 1993; Honkela & Valpola, 2004) allows us to interpret $\mathrm{KL}\left[q(\boldsymbol{\Theta}; \boldsymbol{\Lambda})||p(\boldsymbol{\Theta})\right]$ as the coding length of meta-parameters when encoded together with the data. That means, we effectively control how many bits are required to represent the emerging inference algorithm (i.e., the trained neural network).

The statement that the objective in Equation 7 leads to algorithms that optimally trade-off performance for compressibility is valid for any prior. In this work, we use a sparsity-inducing prior (Kingma et al., 2015; Molchanov et al., 2017; Tipping, 2001), which additionally means that under large resource limitations only networks with few non-zero meta-parameters remain. Thus, resulting algorithms are simple in terms of their statistical complexity and in terms of the number of remaining meta-parameters. Figure 2 schematically contrasts two networks obtained from optimization with low and high resource limitations. In Appendix B we provide further details about the network architecture, training procedure and choice of prior.

Both the single cue heuristic and equal weighting are subsets of the space of all possible weight vectors that can be inferred. Equal weighting heuristics correspond to uniform vectors (e.g., $[1, 1, 1, 1]$), while single cue heuristics can be expressed through a vector with a single non-zero entry (e.g., $[1, 0, 0, 0]$). BMLI could thus – in principle – discover the two heuristics and select between them whenever appropriate.

Note, that BMLI is different from regularizing estimates of regression weights from Equation 2 directly. Regularizing those estimates directly would only slow down the learning progress, but not lead to strategies that are similar to single cue or equal weighting heuristics. Instead, we regularize the meta-parameters of the recurrent neural network generating those weights.

Summary: Computational Models

Let us summarize all outlined models again and contrast the assumptions they make. The ideal observer model assumes that everything about the structure of the decision making environment is known (specifically, it knows about the linear-Gaussian relationship). With this knowledge, it is able to compute the optimal solution by combining information from all features through weighted sums. Heuristics, like the single cue strategy and equal weighting, assume that computing weighted sums is too

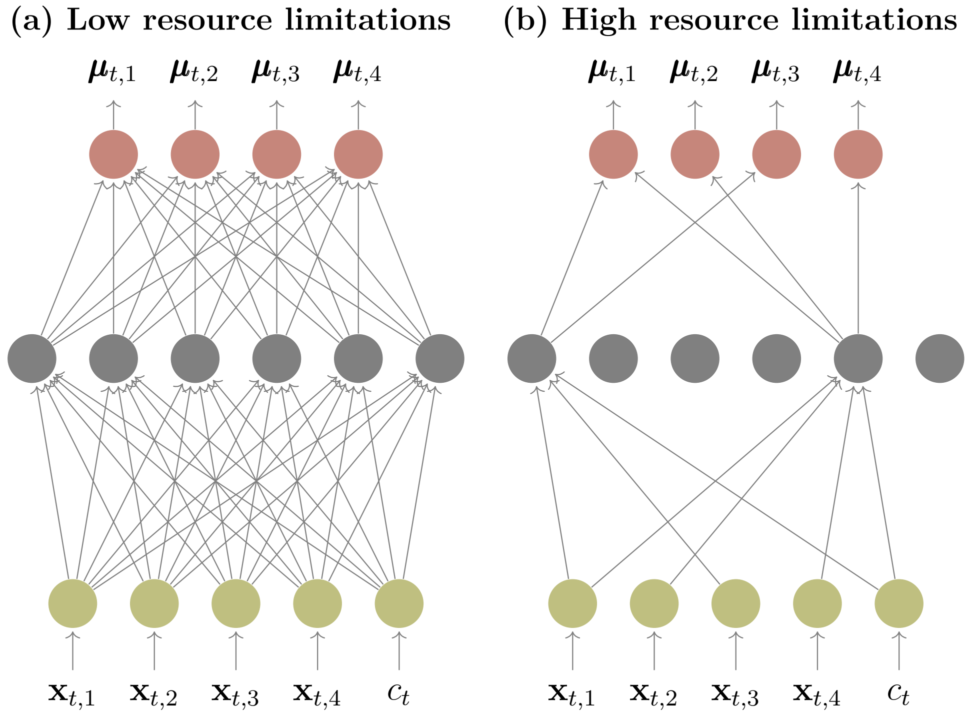

Figure 2

Illustration of two optimized neural networks with a sparsity-inducing prior and different resource limitations. For clarity, we omit recurrent connections and show only the means of $q(\mathbf{w}; \boldsymbol{\lambda}_t)$ *as an output. (a) Network trained with low resource limitations uses all available connections. (b) Network trained with high resource limitations uses only the set of connections that are most useful for increasing performance. Network (b) is much simpler than network (a).*

burdensome and instead bet on simpler ways for making decisions. The single cue heuristic only inspects a single feature, while the equal weighting heuristic sums up all features without weighting them. BMLI does not know anything about the structure of the environment explicitly. Instead, it has acquired a resource rational algorithm to infer decision making strategies through repeated encounters with an environment. Thus, BMLI can exploit characteristics present in that specific environment, while also being efficient in terms of computational resources.

Model Simulations

Next, we demonstrate through a series of model simulations that BMLI recovers both single cue and equal weighting heuristics in specific environments. However, we also identify circumstances where BMLI does not discover any known heuristic and instead infers strategies that use weighted combinations of all features. Before running these simulations, we first have to specify the assumptions we make about the environment and introduce a method for analyzing the emerging strategies.

Environments

For each task in our environment we randomly generate regression weights (Equation 1) by sampling from a standard normal distribution. They are held constant over a task, but vary between tasks. The decision making agent cannot access these weight vectors directly, but instead has to infer them based on observations. All tasks involve two options with four different features and we concentrate on tasks with freely accessible information.

Both redundancy and uncertainty are crucial factors in many real world decision making problems (Gigerenzer & Gaissmaier, 2011). Thus, we want them to be present in our environments. To realize redundancy, we sample features from a multivariate normal distribution with zero mean and covariance matrix $\mathbf{\Sigma}$. Partially redundant features are ensured by drawing separate covariance matrices from a LKJ prior with $\eta = 2$ (Lewandowski et al., 2009) for each task. To introduce uncertainty, decisions are based on only a few examples (between zero and nine) and we set the additive noise term σ in Equation 1 such that an ideal observer with nine prior observations is correct in 85% of the cases.

In contrast to most prior work, we investigate paired comparison tasks with continuous features. In many real world scenarios, features are naturally described through continuous values and thus we believe that the restriction to binary features neglects a characteristic present in many of the problems people typically solve. Moving to continuous features also facilitates statistical analysis as fewer trials are needed to observe expected effects. For example, it would require over four times more trials to distinguish an ideal observer model from the single cue heuristic in environments with dichotomized features instead of continuous ones (see Appendix C for further details).

We consider three variations of the previously outlined environments, that assume (1) known rankings of features, (2) known directions of features or (3) neither. To provide agents with a ranking of features, we arrange them in decreasing order according to the magnitude of their weights. Known directions are ensured by inverting the sign of a feature if it has a negative correlation with the criterion, leading to features with only positive directions.[3] These environments are used for meta-learning and to generate the tasks for our empirical studies.

Gini Coefficients

In order to characterize different decision making strategies, we adopt a measure from the economics literature called the Gini coefficient (Atkinson et al., 1970). The Gini coefficient was originally intended to describe income and wealth distribution of countries. Its minimal value of zero corresponds to a country in which all residents are equally wealthy, while the maximal value of one corresponds to a country in which a single person possesses everything.[4]

The extreme cases of the Gini coefficient also coincide with the two previously discussed heuristics: equal weighting heuristics have a Gini coefficient of zero, while single cue heuristics have a Gini coefficient close to one. Thus, we can employ the Gini coefficient

[3] Our ideal observer implementation always assumes the original standard normal prior over weights, i.e. the prior is not adjusted based on the additional information about ranking or direction.

[4] The extreme value of one is only reached in the limit of an infinite number of residents, otherwise the maximum Gini coefficient for d residents is $1 - d^{-1}$.

to understand how similar estimated regression weights are compared to both heuristics. In practice, we compute Gini coefficients from absolute values of weight vectors. Mathematically, the Gini coefficient of a weight vector $\mathbf{w} \in \mathbb{R}^d$ is defined as half of the relative mean absolute difference:

$$G(\mathbf{w}) = \frac{\sum_{i=1}^{d} \sum_{j=1}^{d} |\mathbf{w}_i - \mathbf{w}_j|}{2d \sum_{i=1}^{d} \mathbf{w}_i} \tag{8}$$

Throughout this section, we analyze Gini coefficients for BMLI (with $\beta = 0.01$), MLI and ideal observer models. If Gini coefficients are consistently close to zero or one, we deduce that the model has recovered one of the two heuristics.

BMLI Discovers Heuristics

First, we consider an environment with known feature rankings. For MLI and BMLI we optimized meta-parameters until convergence in an environment where features are ordered based on the magnitude for their associated weight. We then analyze Gini coefficients of inferred regression weights after meta-learning is completed. Because MLI and BMLI are adapted to the environment, they could exploit the additional ranking information to adjust how they infer strategies.

Figure 3 (a) visualizes Gini coefficients obtained from BMLI. We observe strategies with nearly maximum Gini coefficients, which correspond to weight vectors that only have a single non-zero component. Thus, we conclude that the single cue heuristic emerged as the resource rational strategy for an environment with known feature rankings. Looking at MLI in Figure 3 (b), we find Gini coefficients that cover a much wider range of values. Even though there is an initial tendency towards single cue heuristics, many later decisions are based on compensatory rules. This indicates that being adapted to the environment alone is not a sufficient justification for heuristics. Instead, we need algorithms that are adapted to the environment *and* efficient in terms of their computational resources.

Decisions in the ideal observer model are nearly always based on weighted combinations of multiple features, and hence its Gini coefficients in Figure 3 (c) spread over an even wider range of values.

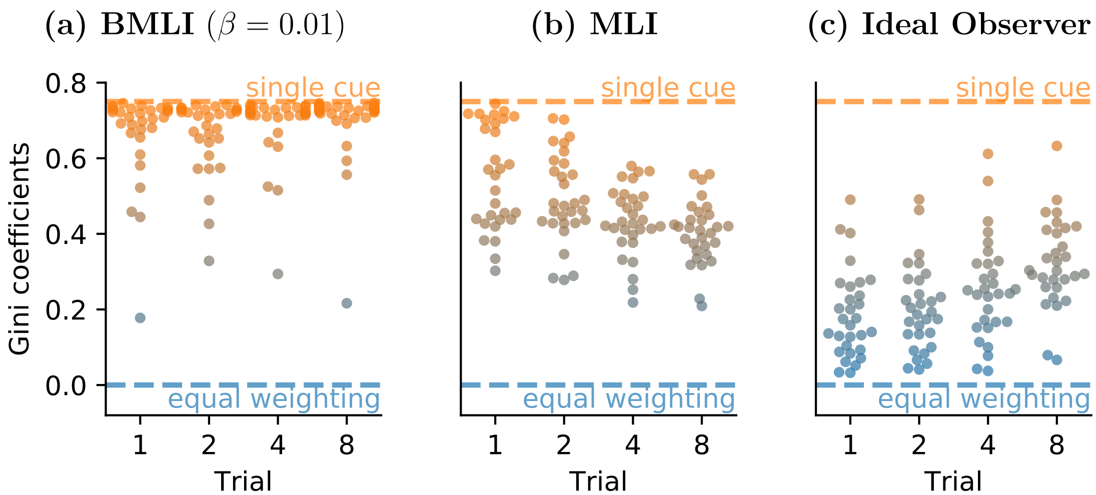

Figure 3

Gini coefficients for an environment with known rankings. High values indicate similarity to the single cue heuristic, while low values correspond to equal weighting heuristics. (a) BMLI results in Gini coefficients that are close to the single cue heuristic. (b) MLI shows tendencies towards the single cue heuristic, especially with few observations. (c) Gini coefficients of the ideal observer model cover the whole range of possible values, indicating that a weighted combination of multiple features is used.

Next, we look at environments where feature directions are known instead of their ranking. For this, we optimize MLI and BMLI in an environment with only positive feature directions. The result here looks very different compared to the ranking condition. Gini coefficients resulting from BMLI, visualized in Figure 4 (a), are consistently close to zero. Low Gini coefficients correspond to uniform weight vectors and hence in this environment the equal weighting heuristic turned out to be rational under limited computational resources. Figure 4 (b) confirms earlier results showing that MLI only leads towards an initial tendency towards heuristics. Early strategies are similar to equal weighting, but especially as more data is observed strategies with higher Gini coefficients emerge. The ideal observer model on the other hand does not exploit environmental characteristics and hence we find no noticeable change in Gini coefficients compared to an environment with known rankings (Figure 4 (c)).

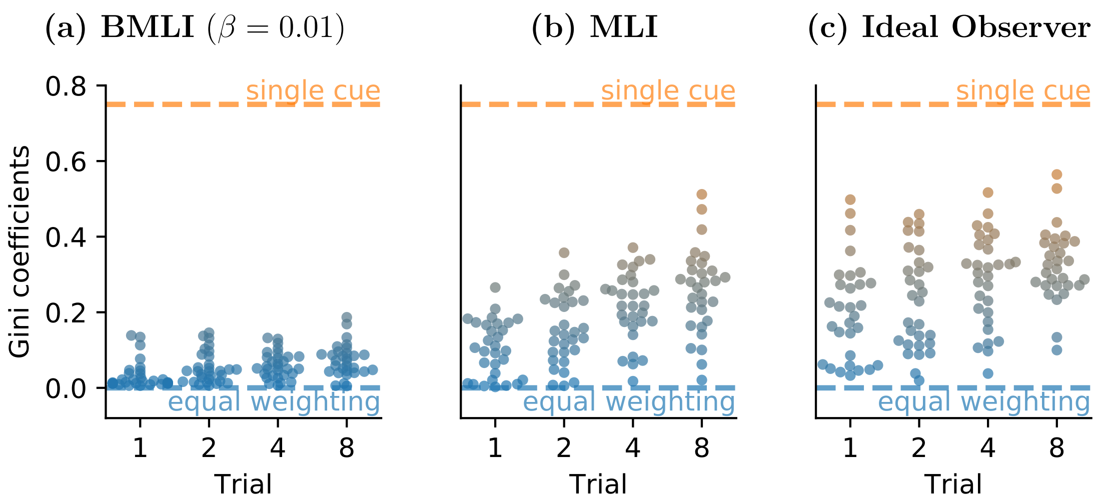

Figure 4

Gini coefficients for an environment with positive directions. High values indicate similarity to the single cue heuristic, while low values correspond to equal weighting heuristics. (a) BMLI results in Gini coefficients that are close to the equal weighting. (b) MLI shows tendencies towards the equal weighting heuristic, especially with few observations. (c) Gini coefficients of the ideal observer model cover the whole range of possible values, indicating that a weighted combination of multiple features is used.

BMLI Does Not Always Discover Heuristics

We have seen that BMLI discovered different heuristics in two classes of environments. Next, we show that there are also environments where this is not the case. For this, we optimized MLI and BMLI such that they adjust to problems without additional information in the form of ranking or direction. Gini coefficients obtained from BMLI reveal that neither single cue nor equal weighting heuristics are resource rational under such circumstances, as shown in Figure 5 (a). Instead, the pattern now looks more similar to one observed in MLI and the ideal observer models, shown in Figures 5 (b) and (c) respectively. In all cases, Gini coefficients cover the full range of possible values, indicating that inferred weight vectors integrate information from multiple features to different degrees.

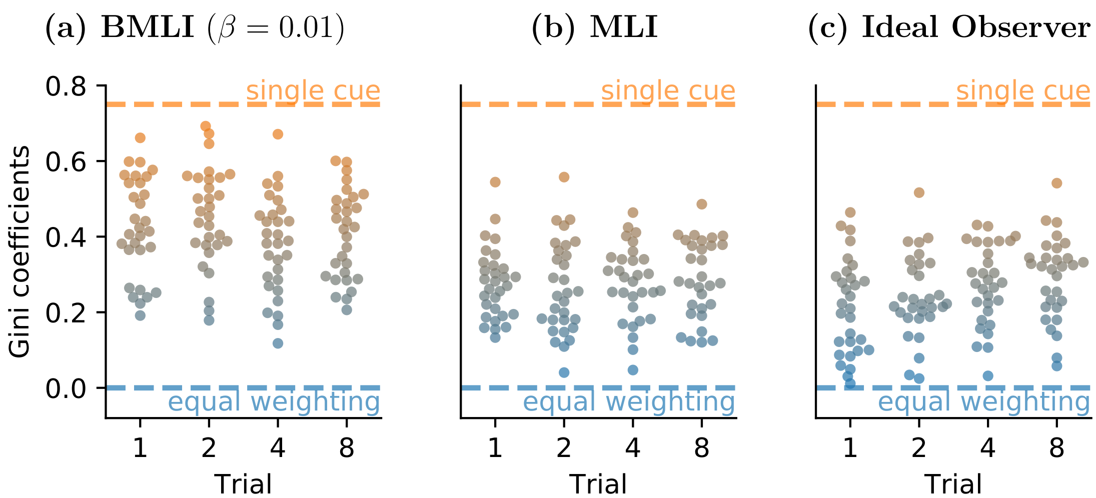

Figure 5

Gini coefficients for an environment without ranking or direction. High values indicate similarity to the single cue heuristic, while low values correspond to equal weighting heuristics. (a) BMLI, (b) MLI and (c) ideal observer models result in Gini coefficients that cover the whole range of possible values, indicating that a weighted combination of multiple features is used.

Experimental Predictions

BMLI discovers both single cue and equal weighting heuristics when information about ranking and direction is provided, respectively. However, resulting strategies diverge from known heuristics whenever such information is not present. Instead, our simulation results suggest that weighted combinations of multiple features should be used in such situations. Under the assumption that people make adaptive and computationally efficient inferences, our results enable us to make precise predictions about when to expect heuristics as part of human decision making and when not: knowing the correct ranking of attributes leads to one reason decision making, knowing the directions of the attributes leads to equal weighting, and not knowing about either leads to strategies that use weighted combinations of multiple attributes. Below, we present results of three paired comparison studies that confirm the predictions made by BMLI.

Experiment 1: Known Ranking

In the first study, participants made decisions in multiple paired comparison tasks while having access to a ranking of features, but not their underlying weights. Previously, we showed that in environments with known feature rankings, single cue heuristics are resource rational strategies. Hence, we hypothesized that people are more likely to apply the single cue heuristic in this condition.

Methods

Participants

Participants were students from the University of Marburg, taking part in the study for course credits. Besides course credits, they got a chance to win a €10 voucher if they made more than 66.6% correct decisions. The experiment was approved by the local ethics board (AZ 2020-32k). In total, we collected data from 28 participants (23 female, average age: 22.36 ± 5.65).

Procedure

Each participant performed 30 different paired comparison tasks that were randomly generated according to the previously described distribution. Each task consisted of ten trials. Underlying weights remained fixed within a task, but varied between tasks. Participants were informed about transitions between tasks. Each participant encountered the same set of paired comparison tasks in a randomized order.

The problem was framed as an alien sports competition on an unknown planet (Figure 6). Participants observed four numerical attributes for two aliens and indicated by a button press which alien they believed was more likely to win. The alien cover story was used to keep the meaning of features completely abstract from the participant's perspective. Participants did not have access to the underlying weights, but instead had to learn about the importance of features based on experience. Feedback about the correct choice was provided directly after each decision. For this condition features were displayed in descending order based on the magnitude of weights. Participants were told that features are arranged from top to bottom according to how well they predict the winner. Being aware about this additional ranking information allowed them to apply strategies that are appropriate for this environment. Participants went through a short tutorial and did a comprehension check to confirm that they understood the instructions. The median time to complete the experiment was 26.00 minutes.

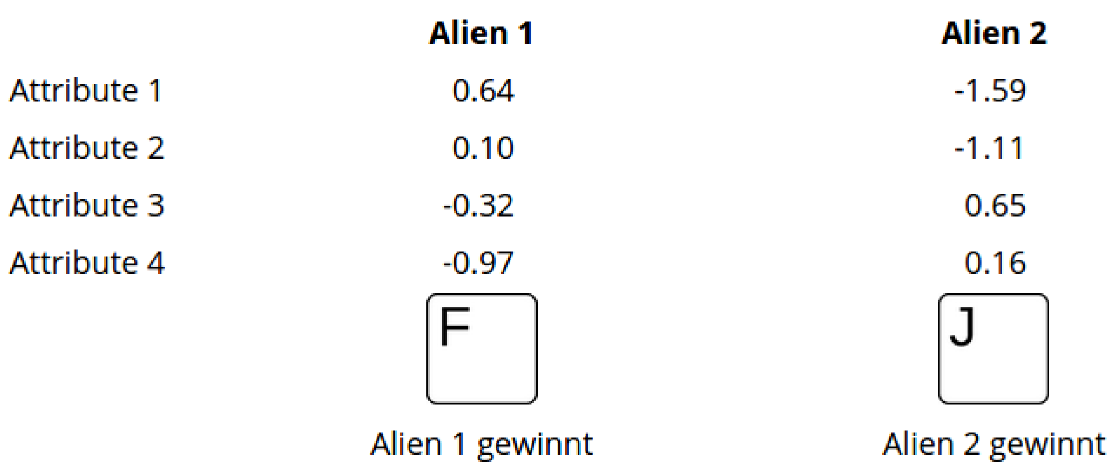

Figure 6

Graphical illustration of a single trial in the experiment. "Alien X gewinnt" translates to "Alien X wins".

Results

Performance

Figure 7 (a) shows the percentage of correct decisions for participants in our study together with the accuracy of different models. Participant performance was within the range of the single cue heuristic and BMLI. On average, participants made $68.25 \pm 7.55\%$ correct choices.

Model Comparison

If people make efficient use of their available computational resources, we expect them to adopt the single cue heuristic in this experiment. To examine this hypothesis, we performed a Bayesian model comparison and computed posterior probabilities of the different models given the decisions made by a participant. Appendix D provides a detailed description of the methods we used for statistical analysis. Because the single cue heuristic and BMLI make redundant predictions, we decided to split our analysis into two parts. First, we analyzed all models except BMLI for individual participants. Then, we compared BMLI against the other models on the data of all participants.

In 22 out of 28 participants, we found evidence for the application of the single cue heuristic. For all of those participants, the model evidence decisively favored the single cue heuristic ($p(m = \text{SC}|\hat{\mathbf{c}}^{(i)}, \mathbf{X}^{(i)}) > 0.99$). Figure 7 (b) summarizes posterior probabilities of different models for all participants. Participants not best described by one reason decision making were instead best described by guessing. The protected exceedance probability (PXP), which measures the probability that a particular model is more frequent in the population than all the other models under consideration (Rigoux et al., 2014), favored the single cue heuristic decisively (PXP > 0.999).

Finally, we compared how well BMLI fared against the other models. Because BMLI also includes guessing with large and compensatory strategies with low resource limitations, it allows us to capture individual differences. For this analysis, we identified

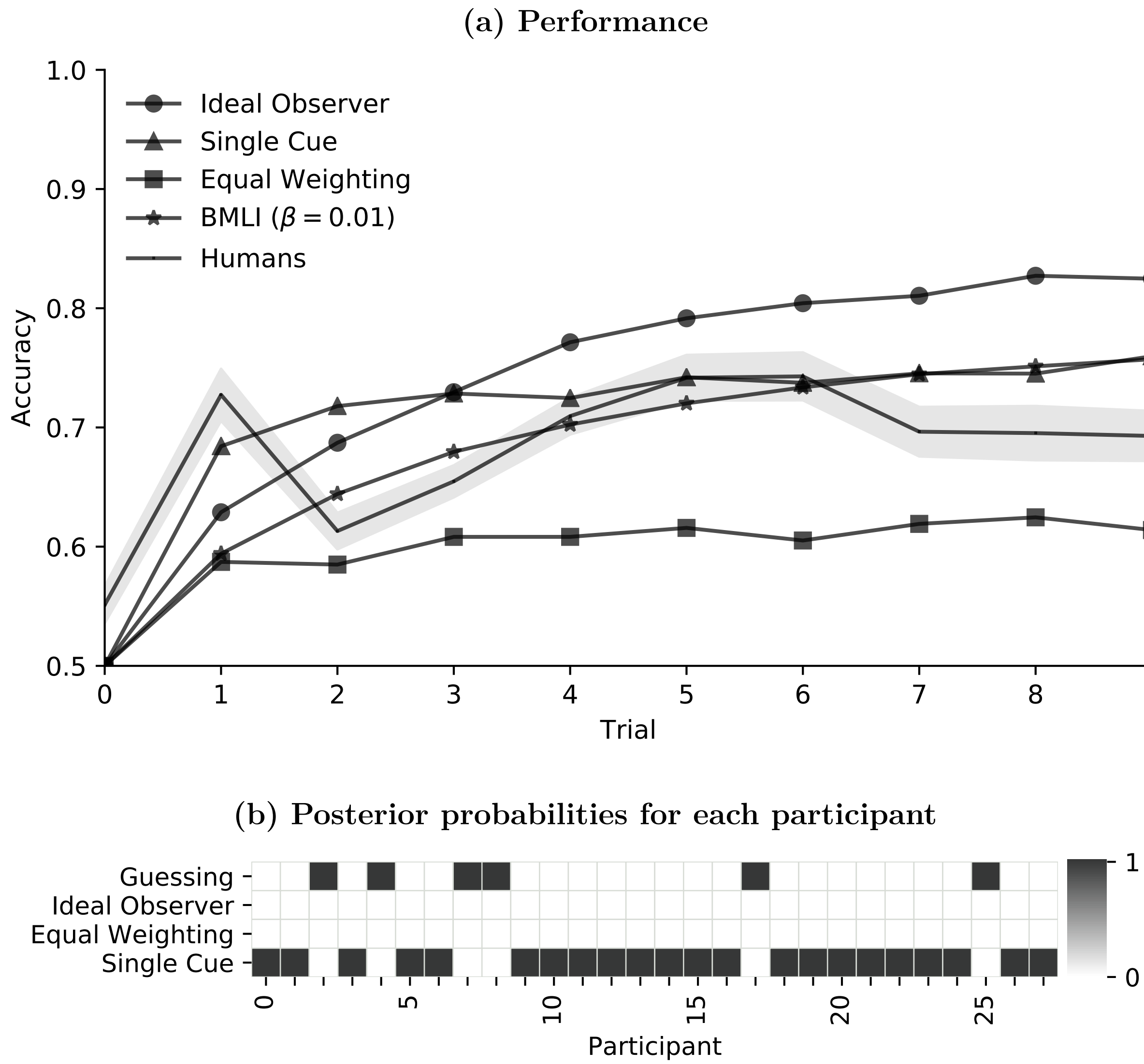

Figure 7

(a) Percentage of correct decisions in the ranking condition plotted over number of observed examples. Shaded contours represent the standard error. (b) Posterior distributions for each participant over different strategies in the ranking condition. High values indicate that the participant was likely to use the corresponding strategy.

the best fitting regularization factor β for each participants and approximate their overall evidence using the Bayesian information criterion (BIC). The resulting posterior probabilities indicated that across all participants BMLI offered an even better explanation for the observed data than the other models ($p(m = \text{BMLI}|\hat{\mathbf{c}}, \mathbf{X}) \approx 1$). This is the case, because BMLI explained the behavior of participants that used single cue heuristics *and* participants that used guessing.

Discussion

Most empirical evidence for one reason decision making has been provided by studies that involved a cost for acquiring information about features (Bröder, 2000; Bröder & Gaissmaier, 2007; Rieskamp & Otto, 2006). However, even with an experimental protocol that favored few pieces of information, evidence for these strategies remained inconclusive (Newell et al., 2003; Scheibehenne et al., 2013). When information is freely available, people are often better described through compensatory strategies such as logistic regression (Bröder, 2000; Glöckner and Betsch, 2008; Lee and Cummins, 2004; Parpart et al., 2018). Our results are among the first to decisively show that people's choices can be based on a single piece of information, even when such strategies are not favored by the experimental protocol. This was possible, because we precisely identified conditions under which one reason decision making *should* appear. Nearly all participants in our study applied strategies that were efficient in terms of resources while also accounting for environmental characteristics.

Experiment 2: Known Direction

In our second study, we provided no information about ranking and instead informed participants about feature directions; otherwise it was identical to the first experiment. In our previous analysis, we have seen that this modification also caused a change in what strategy is resource rational. Now, resource rational decision making amounts to the application of equal weighting heuristics. We therefore hypothesized that participants would become more likely to use such strategies.

Methods

Participants

Participants were students from the University of Marburg, taking part in the study for course credits. Besides course credits they got a chance to win a €10 voucher if they made more than 66.6% correct decisions. The experiment was approved by the local ethics board (AZ 2020-32k). In total, we collected data from 24 participants (22 female, average age: 22.54 ± 3.28).

Procedure

The design was identical to the first experiment, except that participants were informed about the presence of positive feature directions instead of the feature ranking. This was realized by telling them that higher feature values always made it more probable for an alien to win the competition. The median time to complete the experiment was 29.69 minutes.

Results

Performance

Participants made on average $73.85 \pm 4.53\%$ correct choices, putting their performance within the range of all models, see Figure 8 (a). The higher average performance indicates that participants found it overall easier to process information about direction than about ranking. Participants' performance in the initial step turned out to be substantially higher than the ideal observer model and both heuristics, indicating that information about direction is useful even before making observations. This characteristic is also captured in BMLI.

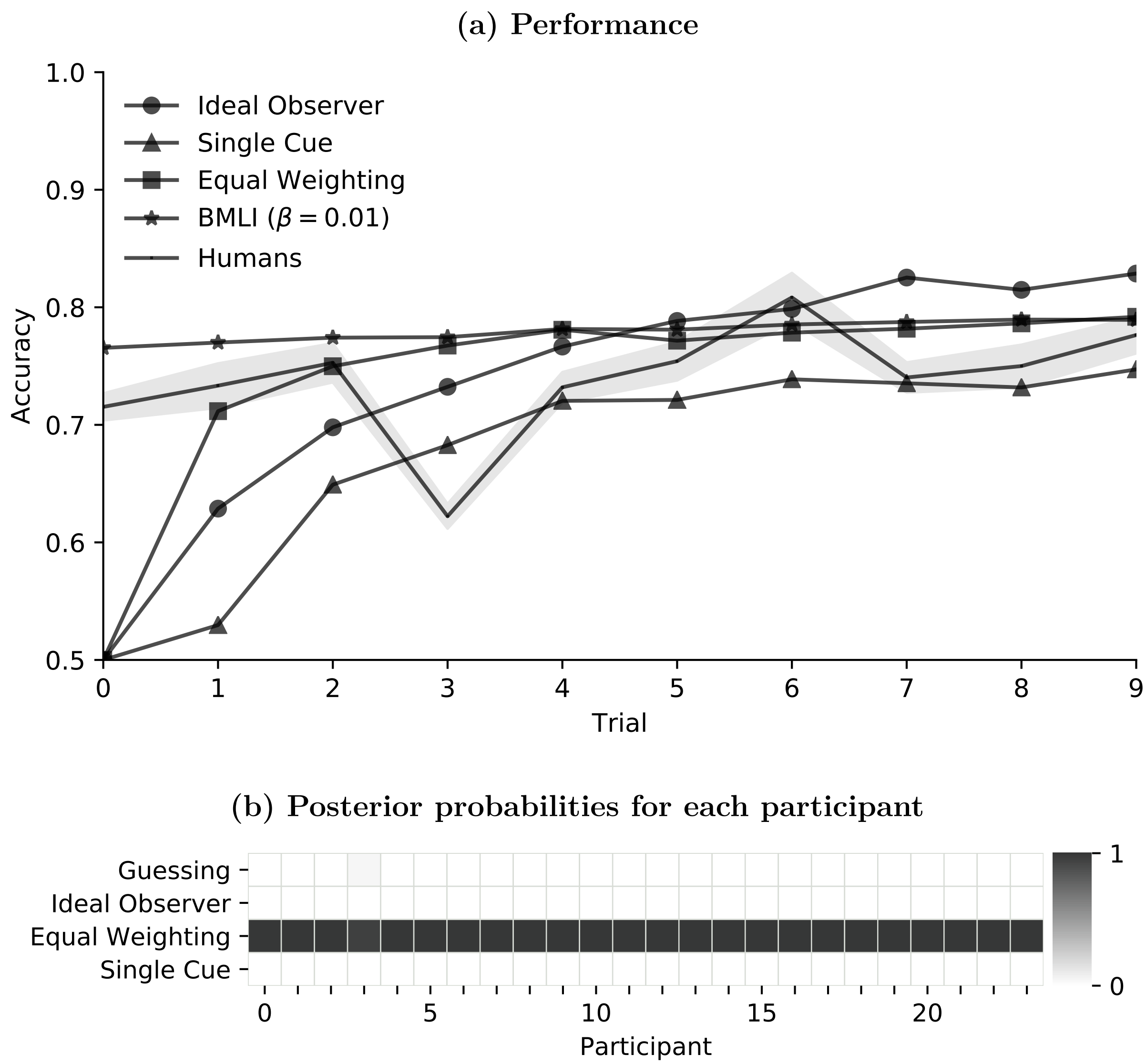

Figure 8

(a) Percentage of correct decisions in the direction condition plotted over number of observed examples. Shaded contours represent the standard error. (b) Posterior distributions for each participant over different strategies in the direction condition. High values indicate that the participant was likely to use the corresponding strategy.

Model Comparison

In this condition, equal weighting and BMLI made partially redundant predictions. Thus, we again decided to split our analysis into two parts. First, we analyzed all models except BMLI for individual participants. Then, we compared BMLI against the other models on the data of all participants.

The posterior probabilities of different models, illustrated in Figure 8 (b), confirmed the prediction of our earlier simulations. Most participants indeed adhered to the resource rational maxim and applied equal weighting heuristics. For all participants, equal weighting provided the best explanation for the observed data. For all but one participant, evidence turned out to be decisive ($p(m = \mathrm{EW}|\hat{\mathbf{c}}^{(i)}, \mathbf{X}^{(i)}) > 0.99$). The probability that equal weighting was the most frequent model in the population (PXP > 0.999) supported the conclusion that people in general applied equal weighting heuristics when information about direction was available.

When additionally comparing BMLI against the other models on the aggregated data of all participants, we found that BMLI again offered an even better explanation than all other models ($p(m = \mathrm{BMLI}|\hat{\mathbf{c}}, \mathbf{X}) \approx 1$). Here, this was the case because BMLI was able to capture participants' decisions in the initial step, while the equal weighting heuristic did not.

Discussion

Similar to the results of our first study, we found that people apply resource rational strategies that are adequate for the given environment. Participants performed better compared to the first study, indicating that they found it easier to work with directions than with rankings. We speculate that one explanation for this observation could be that positive correlations are more frequently encountered in the world.

Previous empirical studies (see our earlier analysis in Tables 1 and 2) on heuristics were often restricted to tasks with positive correlations between features and the criterion.

Despite this, few studies actually consider equal weighting heuristics when comparing their hypotheses. Instead, most of them attempted to show that people rely on one reason decision making, often with inconclusive results. We believe that this mismatch between the hypotheses being tested and the structure of the tasks considered is an important factor in explaining mixed results of prior empirical work.

Experiment 3: Unknown Ranking and Direction

In our final study, we investigated choice behaviour in an environment that did not provide information about ranking or direction. In the previous model simulations, we have demonstrated that no heuristic emerges under such conditions. Instead, BMLI discovered strategies with compensatory weights even under large resource constraints. Hence, we predicted that people in this condition are less reliant on traditional heuristics and instead integrate information from multiple features properly.

Methods

Participants

Participants were students from the University of Marburg, taking part in the study for course credits. Besides course credits, they got a chance to win a €10 voucher if they made more than 60% correct decisions. The experiment was approved by the local ethics board (AZ 2020-32k). In total we collected data from 23 participants (16 female, average age: 23.09 ± 4.38).

Procedure

The design was identical to both previous experiments, except that it did not include information about feature rankings and direction anymore. The median time to complete the experiment was 36.09 minutes.

Results

Performance

The ideal observer model and the equal weighting heuristic remained identical in their performance compared to the first study, see Figure 9 (a). The single cue heuristic however performed slightly worse, as it was not provided with knowledge about the most predictive feature anymore, but instead had to infer it based on observations. Note, that with an identical level of resource limitations the performance of BMLI substantially decreased compared to the previous environments.

Participants also found this version much harder and performed substantially worse. Without the additional information from the first two conditions, their cognitive resource limitations became a dominating factor. The average performance dropped to $57.14 \pm 4.38\%$. While some participants performed well, a substantial amount was at or close to chance level.

Model Comparison

According to our model simulations, we should expect to find evidence for models using weighted combinations of multiple features in this condition. Because no known heuristic emerged in this environment, we did not split our analysis and already considered BMLI on the level of individual participants.

Posterior probabilities obtained from a Bayesian model comparison in Figure 9 (b) confirmed that most participants combined information from multiple features instead of using heuristics like equal weighting or one reason decision making. Fifteen out of 23 participants were best described by BMLI; in nine of those we found decisive evidence ($p(m = \text{BMLI}|\hat{\mathbf{c}}^{(i)}, \mathbf{X}^{(i)}) > 0.99$). We again found that BMLI fared favorably against all other models on the aggregated data ($p(m = \text{BMLI}|\hat{\mathbf{c}}, \mathbf{X}) \approx 1$). The protected exceedance probability (PXP > 0.999) also supported the conclusion that BMLI was the most frequent explanation for participants in our population.

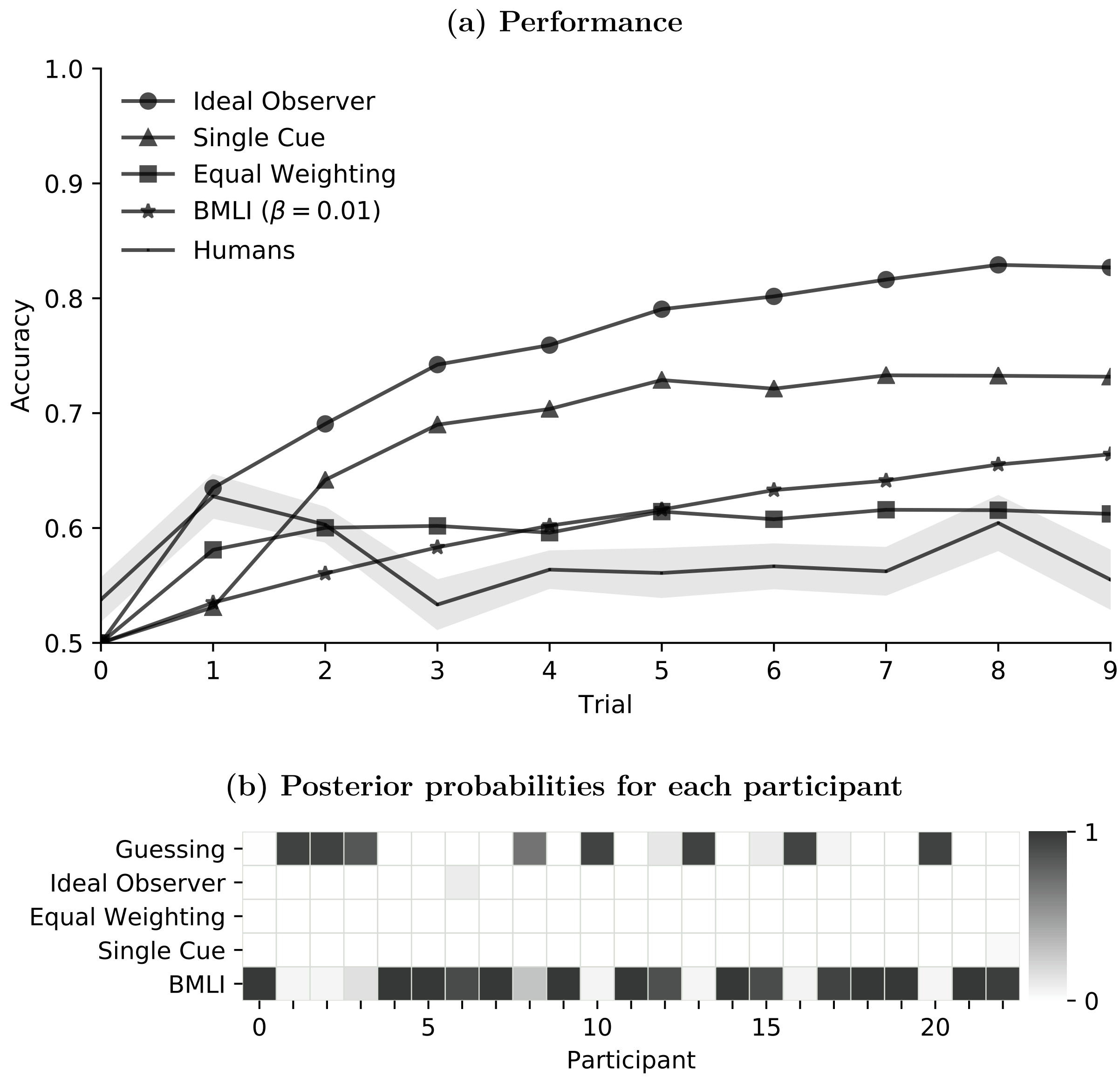

Figure 9

(a) Percentage of correct decisions in the unrestricted condition plotted over number of observed examples. Shaded contours represent the standard error. (b) Posterior distributions for each participant over different strategies in the unrestricted condition. High values indicate that the participant was likely to use the corresponding strategy.

Discussion

In an environment that did not provide additional information about ranking or direction, participants' decision making again followed the prediction made by BMLI. Most participants applied strategies that involved weighted combinations of features, as it was suggested by our model simulations. The general result that most people were able to quickly combine information from multiple sources if needed is also consistent with results of prior studies (Bröder, 2000; Glöckner & Betsch, 2008; Parpart et al., 2018).

General Discussion

At the core of theories of ecological rationality, researchers have posited an interaction between cognition and the environment. Brunswik (1956) argued that human perception cannot be understood in laboratory settings alone, but rather has to be interpreted in the light of real environments in which real objects are perceived and acted upon. Simon (1990b) famously highlighted the interaction between cognition and the environment using an analogy of a pair of scissors, with one blade being the structure of the environment and the other blade the computational capabilities of the subject. This conceptualization of ecological rationality has strongly influenced theories of heuristic decision making. The need to economize cognitive resources places pressure on the mind to employ heuristics that work well in specific environments. Nonetheless, how people pick a particular heuristic for a specific environment and where those heuristics come from in the first place has remained elusive. The theoretical picture becomes even more puzzling when looking at the empirical support for heuristic decision making. Proponents of heuristic decision making acknowledge these problems. For example, Gigerenzer (2008) writes: "Why do heuristics work? They exploit evolved capacities that come for free. In addition, they are tools that have been customized to solve diverse problems. By understanding the ecological rationality of a heuristic, we can predict when it fails and succeeds. The systematic study of the environments in which heuristics work is a fascinating topic and is still in its infancy.” But what does a theory, which can explain how heuristics emerge and how they are selected while at the same time accounting for the sometimes mixed empirical results, look like?

We have put forward BMLI as theory that makes significant advances on these questions. Our simulation results show that BMLI discovers previously suggested heuristics. Thus, it provides a normative justification for heuristic decision making. Moreover, we find that different heuristics emerge depending on environmental assumptions. Thus, BMLI also explains how decision making strategies are selected. Finally, our account generates predictions about if and when a specific heuristic should be applied. Since we find that one reason decision making is unlikely to occur in many of the past experimental set-ups, this explains the mixed results of prior empirical work.

Already early on, researchers working on heuristic decision making levied the criticism that simply observing behavioral biases is not enough, and that “in place of plausible heuristics that explain everything and nothing – not even the conditions that trigger one heuristic rather than another – we need models that make surprising (and falsifiable) predictions” (Gigerenzer, 1996). However, the very fact that several heuristic components have been claimed to be part of a heuristic toolbox without fully specifying how they are selected and combined, has subjected heuristic theories to a similar line of criticism: “If one cannot predict which heuristics will be used in which environments, determining the heuristic that will be selected from the toolbox for a particular environment, the approach looks dangerously unfalsifiable” (Newell et al., 2003). In contrast to these arguments, BMLI makes clear, falsifiable and surprising predictions about when people should apply which heuristic. Specifically, our simulation results show that there are three important classes of environments triggering three decision making strategies. If people know the correct ranking of attributes but not their weights, then they should exhibit one reason decision making. If people know the direction of the attributes but not their ranking, then they should exhibit equal weighting strategies. Finally, if people do not know either the ranking or the direction of the attributes, then they should exhibit strategies that use weighted combinations of attributes. We subjected these predictions to a rigorous test in three paired comparison experiments.

We found that the vast majority of participants applied decision making strategies as predicted by BMLI. Moreover, BMLI captured elements of human decision making that could not be explained by traditional heuristics in all three experiments: In the first study it additionally accounted for participants that resorted to guessing, in the second study it provided an explanation for the good initial performance of participants and in the third study it predicted correctly that performance should decrease and that people apply compensatory strategies instead of established heuristics. These results enrich our theoretical and empirical understanding of ecologically rational decision making.

Limitations

Gigerenzer and Todd (1999) argue that decision making under limited resource cannot be expressed through models that perform optimization under constraints: "Optimization under constraints also limits search, but does so by computing the optimal stopping point, that is, when the costs of further search exceed the benefits." Computing this optimal stopping point can be at least as expensive as finding the optimal solution; hence it defeats the initial intention of modeling decision making under resource limitations (Gigerenzer & Todd, 1999; Scheibehenne & Von Helversen, 2009). BMLI involves optimization under constraints, but importantly does so at the meta-learning level, which happens on a much larger time scale (e.g. through evolutionary processes). Learning within an individual task on the hand is fast as it does not involve any form of optimization. This perspective of learning at multiple scales is also at the core of recent theories of fast and slow reinforcement learning (Botvinick et al., 2019).

BMLI assumes that meta-learning happened prior to the experiment, but it remains agnostic about the exact processes controlling the acquisition of strategies. BMLI could, for example, be acquired through evolutionary processes, through individual experiences, or both. If meta-learning indeed happened prior to the experiment, we should find no noticeable improvement in performance over the course of our studies. We find support for this hypothesis when comparing human performance in the first and second half of our studies (Figure 10). Furthermore, we evaluated posterior probabilities of different models for each task as opposed to for each individual participant (Figure 10) and found that participants did not apply different strategies during the experiment. Nonetheless, a valid criticism of our current work is that it does not address the precise process of meta-learning, and whether this process is rather shaped by ontogeny, phylogeny, or both. This is indeed an open problem for all theories of heuristic decision making, which at various times have argued that heuristics emerge from evolutionary pressures (Hutchinson & Gigerenzer, 2005), developmental processes (Gigerenzer, 2003), or task-specific adaptations (Marewski & Schooler, 2011). The time scale of meta-learning therefore remains an open theoretical and empirical question.

We have used a particular model architecture to simulate behavior in our tasks. In particular, we applied a gated recurrent network and adapted the meta-parameters through gradient descent on a loss function that can trade-off between the accuracy of the network and the statistical complexity of its parameters. Thus, a naturally arising question is how much our results depend on the chosen architecture. For the sake of the resource rational argument, we should have used the architecture that optimally solves the accuracy-effort trade-off. Because identifying this architecture is not possible, we settled for the next best option and used an architecture that is known to work well across a wide range of domains. Theoretically, a resource rational algorithm should also be able to recover optimal decision making if there are no resource limitations. Infinitely wide recurrent neural networks are known to be Turing-complete and hence are in theory able to implement optimal decision making (Siegelmann & Sontag, 1992). We confirmed that our networks are wide enough to closely approximate the ideal observer model (Figure 11).

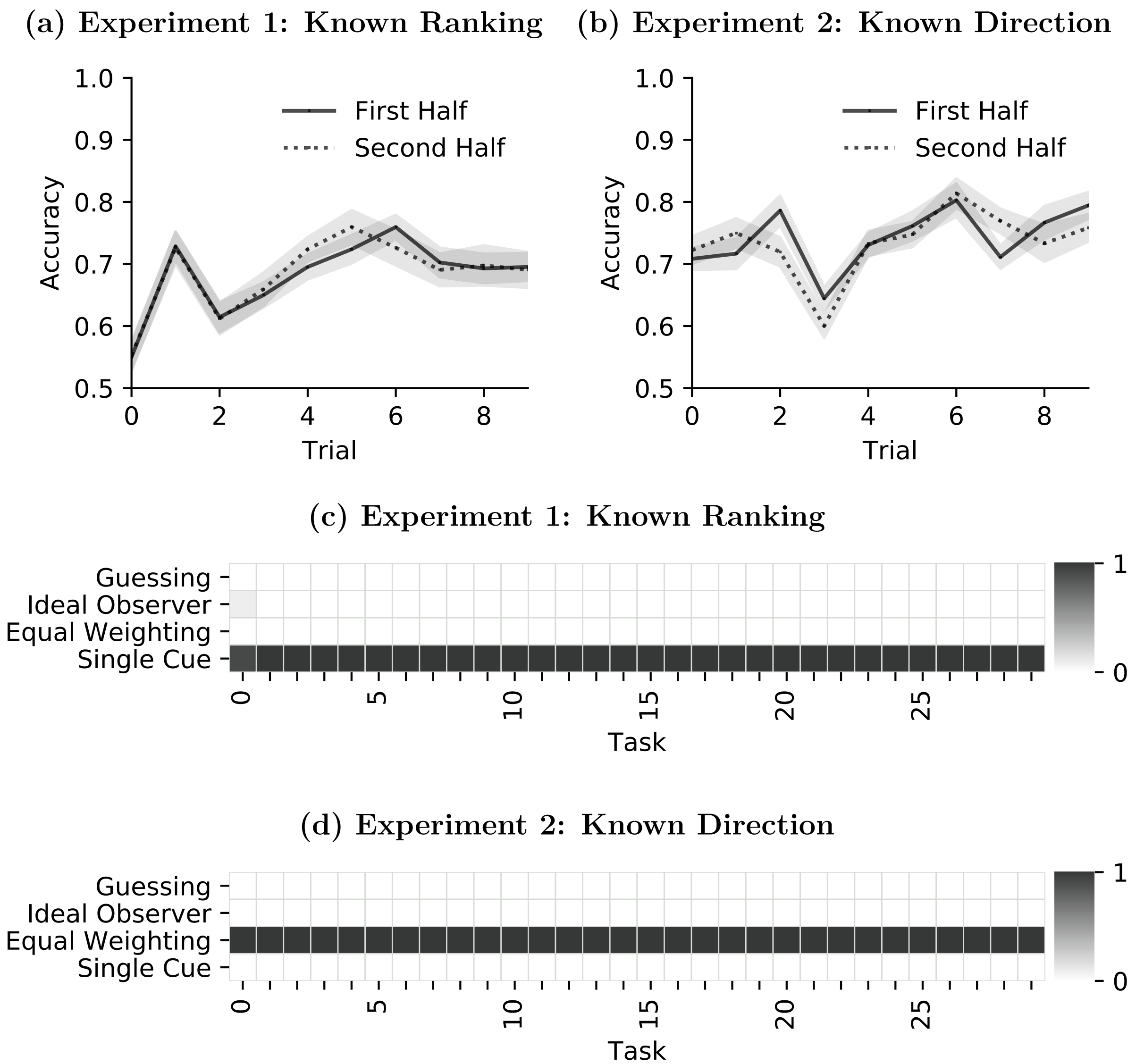

Figure 10

(a) and (b) show that performance of participants did not change over the experiment, indicating that meta-learning already happened prior to the experiment. Shaded contours represent the standard error. (c) and (d) confirm this observation by showing that the selection of strategies also did not change during the experiment. High values indicate that the corresponding strategy was applied with high probability in the given task.

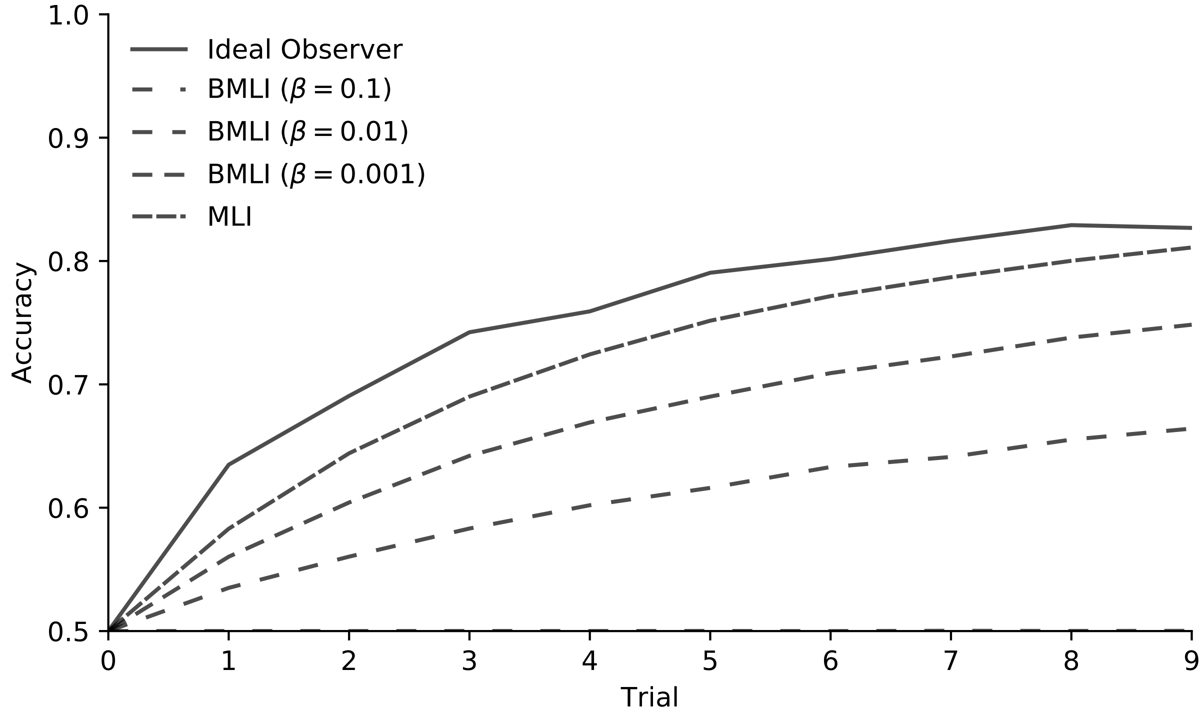

Figure 11

Percentage of correct decisions in an environment without information about ranking and direction plotted over number of observed examples. We observe that the MLI closely approximates the ideal observer, verifying that our models are in general able to implement optimal decision making. Note, that BMLI with $\beta = 0.1$ performs at chance level, meaning that guessing is recovered in the other extreme case of very large resource limitations.

Currently, our approach also does not directly offer a way to predict which properties of the environment will determine what type of decision making strategies are ecologically rational. Instead, we have to train our meta-learning models in different environments and then analyze what decision strategies emerge, for example by analyzing the weights' Gini coefficient. Looking at a model's emerging properties is a common method when neural network approaches are applied to psychological questions (Ritter et al., 2017). We believe that this possible weakness can also be a strength, because it forces researchers to truly study the properties of environments, as has been the core proposal of theories of ecological rationality for decades.

Related Work

To highlight what BMLI adds to existing theories, we compare it to other ideas put forward in previous investigations. In the context of decision making, we focus on methods that address how strategies are selected and how they are discovered. Beyond that, we discuss how meta-learning and resource rationality have been applied to understand other phenomena.

Strategy Selection

First, there have been several theories explaining how strategies are selected. Rieskamp and Otto (2006) proposed a theory of strategy selection learning that framed the strategy selection process as a model-free reinforcement learning problem. Their theory assumes that people slowly learn how to select the right strategy from a given repertoire of strategies based on repeated interactions. A key finding of their experiments was that over time participants learned to select the best-performing strategy for a particular environment. Their method requires learning from scratch whenever it encounters novel problems and hence it does not address how knowledge is transferred between different environments, and why participants are immediately able to select appropriate strategies in our experiments.

Lieder and Griffiths (2017) addressed the missing ability to transfer knowledge between environments through an approach based on rational meta-reasoning. Based on properties of the environment, they predicted speed and accuracy of different strategies. They showed that participants selected the strategy that was best for solving the speed-accuracy trade-off in the current context. In contrast to their work, we used separate models for each environment. However, it would be possible to extend our modeling framework by conditioning the initial state of the recurrent network on features of an environment.

Marewski and Schooler (2011) postulated a probability landscape describing an individual's ability to apply a strategy as a function of cognitive capabilities and the environment. Their work referred to situations in which a strategy can be applied as a cognitive niche and showed that cognitive niches of different strategies are disjoint in many cases. This greatly simplified the the strategy selection problem and was in line with participants' behavior across a number of experiments. We believe that cognitive niches could also be the result of meta-learning, where an algorithm adapts to a given characteristic of environment until it cannot easily be applied to a vastly different environment anymore.

Previous theories of strategy selection all require to define a set of potential strategies in advance. In contrast, BMLI is not restricted to predefined sets and instead discovers useful strategies on the fly.

Strategy Discovery

There have also been some accounts that explain how strategies are discovered. Schulz et al. (2016) proposed a method for learning decision making strategies from small, probabilistic building blocks. Based on a self-reinforcing sampling scheme, they were able to build tree-like non-compensatory heuristics. Their approach can recover TTB on data sets that have been generated by the TTB heuristic. However, it is not able to learn about other, non-compensatory strategies or to make predictions about when participants would prefer which strategy.

Lieder et al. (2017) suggested a model that composes strategies from atomic computations. According to their theory, an agent represents computations as costly actions in a meta-level Markov decision process. The agent's goal is to maximize the external payoff obtained from making correct decisions while accounting for the computational costs of actions. When they applied their theory to several decision making problems, they found that it discovered two known heuristics – TTB and guessing – as well as a novel strategy that combined TTB with satisficing (Simon, 1956).

Parpart et al. (2018) showed that heuristics can emerge from Bayesian inference in the limit of infinitely strong priors. Using this idea, they identified priors corresponding to an equal weighting heuristic. Finding a prior that leads to TTB proved to be more challenging in the Bayesian framework and was only possible after introducing an additional decision rule. Instead of relying on the complexity argument as justification for heuristics, their analysis suggested that heuristics work well because they implement priors that reflect the actual structure of the environment.

Theories that build algorithms from simpler computations (Lieder et al., 2017; Schulz et al., 2016) discover one reason decision making heuristics without difficulties, but struggle to account for equal weighting heuristics. Theories based on Bayesian inference (Parpart et al., 2018) on the other hand have no difficulties with discovering equal weighting heuristics, but require additional components to find heuristics that rely on a single piece of information. We show that people actually use both classes of strategies and provide a theory that can discover both of them in an appropriate context. While there exist prior approaches that address either the strategy selection problem or the strategy discovery problem independently, BMLI is also the first to account for both problems jointly within a unified framework.

Resource Rationality

Achille and Soatto (2018) used an objective identical to Equation 7 to investigate theoretical properties of neural networks outside of the meta-learning setting. Other information-theoretic objectives have been used to study resource rational behavior in a number of contexts (Gershman, 2020; Ho et al., 2020; Ortega & Braun, 2013; Zaslavsky et al., 2018). However, none of these objectives has been applied to limit the computational complexity of meta-learned algorithms.

More relevant to our work is the approach of Dasgupta et al. (2020), who taught neural networks to approximate Bayesian inference, given some information about an inference problem's prior and likelihood. Restricting the size of the network allowed them to account for a large amount of cognitive biases, including base rate neglect and conservatism. This approach shares its core principles with our theory: resource rationality and meta-learning. However, BMLI does not approximate Bayesian inference explicitly as done by Dasgupta et al. (2020). Instead, it attempts to infer distributions that are optimal for making future predictions (which may or may not correspond to Bayesian inference).

Meta-Learning as Theory of Human Behavior

Brighton (2006) and Chater et al. (2003) considered standard feed-forward networks trained with backpropagation as models of decision making in paired comparison tasks. Their results indicated that, if only a few examples were used, such models tended to overfit and were outperformed by much simpler, more robust alternatives. Brighton (2006) suggested meta-learning as a potential solution to this problem of overfitting, but did not provide a concrete implementation of this conjecture. BMLI is such an implementation that can be applied to paired comparison tasks with few examples and – crucially – *without* showing signs of overfitting. Key to BMLI's success is that learning happens solely in the fully-trained network's recurrent activations and not through traditional gradient-based training schemes.

When we look beyond decision making and paired comparison tasks, meta-learning has recently received increased attention as an explanation for human behavior across a variety of cognitive and neuroscientific questions. For example, meta-learning has been shown to lead to human-like characteristics in the contexts of few-shot learning (Santoro et al., 2016), systematic compositionality (Lake, 2019), exploration (Binz & Endres, 2019) as well as one-shot navigation and model-based reasoning (Wang et al., 2016). The current work adds heuristic strategies of decision making as another domain to this list.

Future Directions

Most computational models in psychology and cognitive science are confined to idealized settings. BMLI on the other hand can – in principle – scale to much more complex domains (Santoro et al., 2016; Wang et al., 2016). Having access to such models allows us to study human behavior under more realistic conditions. In the context of decision making it becomes, for example, possible to investigate how and why different representational formats influence human strategies (Bröder & Schiffer, 2006) by learning models that directly process visual representations of the task.

In this paper we have applied BMLI to the paired comparison setting. However, BMLI is more general than that and we believe that it could also be used to explain heuristics in other context, such as the recognition heuristic (Goldstein & Gigerenzer, 2002) or the gaze heuristic (Belousov et al., 2016; Shaffer et al., 2004). BMLI could also provide insights into other phenomena in human learning, such as the observation that learning about multiple tasks is usually easier when tasks are presented successively compared to an interleaved presentation (Flesch et al., 2018).

The classical approach to computational modeling is to propose a model, test its predictions and finally revise the model if required. However, we can also envision an approach for the revision of theories that puts the study of environments first. In this framework we would ask ourselves what environments can account for observed behavior assuming that people make ecologically and resource rational decisions, instead of revising arbitrary parts of the model. That this is a promising research direction for building more human-like agents was shown for example by Hill et al. (2020), who demonstrated that systematic generalization can be an emergent property of an agent interacting with a *rich* environment.

Finally, our theory provides us with a set of predictions about what should happen when available computational resources are manipulated. It will be interesting to see whether people follow the behavioral trajectories stipulated by BMLI when put under cognitive load or whether patients with attention or memory impairment are better described by models with lower complexity.

Conclusion

The idea that theories of human cognition should consider both the structure of the environment and the computational capabilities of the subject has been a central theme in psychology (Simon, 1990b; Todd & Gigerenzer, 2012). However, actual implementations of this principle have been lacking so far. BMLI provides such an implementation by combining the ideas of resource rationality and meta-learning. BMLI accounts for two open questions in the decision making literature simultaneously, explaining why different strategies emerge and how appropriate strategies are selected. By mapping out environments that cause different strategies to be resource rational, we obtain precise predictions about when previously suggested heuristics should be used and when not. We confirmed these predictions in three paired comparison experiments. Taken together, BMLI offers a normative and empirically supported theory of human decision making.

References

Achille, A., & Soatto, S. (2018). Emergence of invariance and disentanglement in deep representations. *The Journal of Machine Learning Research*, *19*(1), 1947–1980.

Achille, A., & Soatto, S. (2019). Where is the information in a deep neural network? *arXiv preprint arXiv:1905.12213*.

Atkinson, A. B. Et al. (1970). On the measurement of inequality. *Journal of economic theory*, *2*(3), 244–263.

Ayal, S., & Hochman, G. (2009). Ignorance or integration: The cognitive processes underlying choice behavior. *Journal of Behavioral Decision Making*, *22*(4), 455–474.

Belousov, B., Neumann, G., Rothkopf, C. A., & Peters, J. R. (2016). Catching heuristics are optimal control policies, In *Advances in neural information processing systems*.

Bengio, Y., Bengio, S., & Cloutier, J. (1991). Learning a synaptic learning rule, In *Ijcnn-91-seattle international joint conference on neural networks*. IEEE.

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 107.

Binz, M., & Endres, D. (2019). Where do heuristics come from? CogSci.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends in cognitive sciences*, *23*(5), 408–422.

Brighton, H. (2006). Robust inference with simple cognitive models., In *Aaai spring symposium: Between a rock and a hard place: Cognitive science principles meet ai-hard problems*.

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

Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review*, *14*(5), 895–900.

Bröder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: Probabilistic inferences from memory and effects of reprensentation format. *Journal of Experimental Psychology: General*, *132*(2), 277.

Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal strategy selection. *Journal of Behavioral Decision Making*, *19*(4), 361–380.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Univ of California Press.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of economic Literature*, *43*(1), 9–64.

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.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics? In *Simple heuristics that make us smart*. Oxford University Press.

Dasgupta, I., Schulz, E., Tenenbaum, J. B., & Gershman, S. J. (2020). A theory of learning to infer. *Psychological Review*, *127*(3), 412.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological bulletin*, *81*(2), 95.

Dieckmann, A., & Rieskamp, J. (2007). The influence of information redundancy on probabilistic inferences. *Memory & Cognition*, *35*(7), 1801–1813.

Einhorn, H. J., & Hogarth, R. M. (1975). Unit weighting schemes for decision making. *Organizational behavior and human performance*, *13*(2), 171–192.

Flesch, T., Balaguer, J., Dekker, R., Nili, H., & Summerfield, C. (2018). Comparing continual task learning in minds and machines. *Proceedings of the National Academy of Sciences*, *115*(44), E10313–E10322.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological review*, *96*(2), 267.

Gershman, S. J. (2020). Origin of perseveration in the trade-off between reward and complexity. *Cognition*.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to kahneman and tversky.

Gigerenzer, G. (2003). The adaptive toolbox and lifespan development: Common questions? In *Understanding human development*. Springer.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on psychological science*, *3*(1), 20–29.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451–482.

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: The take the best heuristic, In *Simple heuristics that make us smart*. Oxford University Press.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of experimental psychology: Learning, memory, and cognition*, *34*(5), 1055.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological review*, *109*(1), 75.

Grünwald, P. D., & Grunwald, A. (2007). *The minimum description length principle*. MIT press.

Hilbig, B. E. (2010). Reconsidering "evidence" for fast-and-frugal heuristics. *Psychonomic Bulletin & Review*, *17*(6), 923–930.

Hill, F., Lampinen, A., Schneider, R., Clark, S., Botvinick, M., McClelland, J. L., & Santoro, A. (2020). Environmental drivers of systematicity and generalization in a situated agent, In *International conference on learning representations*. https://openreview.net/forum?id=SklGryBtwr

Hinton, G. E., & Van Camp, D. (1993). Keeping the neural networks simple by minimizing the description length of the weights, In *Proceedings of the sixth annual conference on computational learning theory*.

Ho, M. K., Abel, D., Cohen, J. D., Littman, M. L., & Griffiths, T. L. (2020). The efficiency of human cognition reflects planned information processing, In *Proceedings of the 34th aaai conference on artificial intelligence*.

Honkela, A., & Valpola, H. (2004). Variational learning and bits-back coding: An information-theoretic view to bayesian learning. *IEEE transactions on Neural Networks*, *15*(4), 800–810.

Hutchinson, J. M., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural processes*, *69*(2), 97–124.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., & Saul, L. K. (1999). An introduction to variational methods for graphical models. *Machine learning*, *37*(2), 183–233.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the american statistical association*, *90*(430), 773–795.

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Kingma, D. P., Salimans, T., & Welling, M. (2015). Variational dropout and the local reparameterization trick, In *Advances in neural information processing systems*.

Lake, B. M. (2019). Compositional generalization through meta sequence-to-sequence learning, In *Advances in neural information processing systems*.

Lee, M. D., & Cummins, T. D. (2004). Evidence accumulation in decision making: Unifying the "take the best" and the "rational" models. *Psychonomic bulletin & review*, *11*(2), 343–352.

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of multivariate analysis*, *100*(9), 1989–2001.

Lichtenberg, J. M., & Şimşek, Ö. (2017). Simple regression models, In *Imperfect decision makers: Admitting real-world rationality*.

Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, *124*(6), 762.

Lieder, F., & Griffiths, T. L. (2019). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 1–85.

Lieder, F., Krueger, P. M., & Griffiths, T. (2017). An automatic method for discovering rational heuristics for risky choice., In *Cogsci*.

Luan, S., Schooler, L. J., & Gigerenzer, G. (2014). From perception to preference and on to inference: An approach–avoidance analysis of thresholds. *Psychological Review*, *121*(3), 501.

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

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological review*, *118*(3), 393.

Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, *52*, 29–71. https://doi.org/10.1023/A:1015516217425

Molchanov, D., Ashukha, A., & Vetrov, D. (2017). Variational dropout sparsifies deep neural networks, In *Proceedings of the 34th international conference on machine learning-volume 70*. JMLR. org.

Newell, B. R., & Lee, M. D. (2011). The right tool for the job? comparing an evidence accumulation and a naive strategy selection model of decision making. *Journal of Behavioral Decision Making*, *24*(5), 456–481.

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.

Ortega, P. A., & Braun, D. A. (2013). Thermodynamics as a theory of decision-making with information-processing costs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2153), 20120683.

Parpart, P., Jones, M., & Love, B. C. (2018). Heuristics as bayesian inference under extreme priors. *Cognitive psychology*, *102*, 127–144.

Parpart, P., Schulz, E., Speekenbrink, M., & Love, B. C. (2017). Active learning reveals underlying decision strategies. *bioRxiv*, https://www.biorxiv.org/content/early/2017/12/25/239558.full.pdf. https://doi.org/10.1101/239558

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of experimental psychology: Learning, Memory, and Cognition*, *14*(3), 534.

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

Reddi, S. J., Kale, S., & Kumar, S. (2019). On the convergence of adam and beyond. *arXiv preprint arXiv:1904.09237*.

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell?

Rieskamp, J., & Otto, P. E. (2006). Ssl: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207.

Rigoux, L., Stephan, K. E., Friston, K. J., & Daunizeau, J. (2014). Bayesian model selection for group studies—revisited. *Neuroimage*, *84*, 971–985.

Ritter, S., Barrett, D. G., Santoro, A., & Botvinick, M. M. (2017). Cognitive psychology for deep neural networks: A shape bias case study, In *Proceedings of the 34th international conference on machine learning-volume 70*. JMLR. org.

Samuels, R., Stich, S., & Bishop, M. (2012). Ending the rationality wars. *Collected Papers, Volume 2: Knowledge, Rationality, and Morality, 1978-2010*, *2*, 191.

Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., & Lillicrap, T. (2016). Meta-learning with memory-augmented neural networks, In *International conference on machine learning*.

Scheibehenne, B., Rieskamp, J., & Wagenmakers, E.-J. (2013). Testing adaptive toolbox models: A bayesian hierarchical approach. *Psychological review*, *120*(1), 39.

Scheibehenne, B., & Von Helversen, B. (2009). Useful heuristics, In *Making essential choices with scant information*. Springer.

Schmidhuber, J., Zhao, J., & Wiering, M. (1996). Simple principles of metalearning. *Technical report IDSIA*, *69*, 1–23.

Schulz, E., Speekenbrink, M., & Meder, B. (2016). Simple trees in complex forests: Growing take the best by approximate bayesian computation. Cognitive Science Society.

Schwarz, G. Et al. (1978). Estimating the dimension of a model. *The annals of statistics*, *6*(2), 461–464.

Shaffer, D. M., Krauchunas, S. M., Eddy, M., & McBeath, M. K. (2004). How dogs navigate to catch frisbees. *Psychological Science*, *15*(7), 437–441.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological bulletin*, *134*(2), 207.

Siegelmann, H. T., & Sontag, E. D. (1992). On the computational power of neural nets, In *Proceedings of the fifth annual workshop on computational learning theory*.

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

Simon, H. A. (1990a). Bounded rationality, In *Utility and probability*. Springer.

Simon, H. A. (1990b). Invariants of human behavior. *Annual review of psychology*, *41*(1), 1–20.

Thrun, S., & Pratt, L. (1998). Learning to learn: Introduction and overview, In *Learning to learn*. Springer.

Tipping, M. E. (2001). Sparse bayesian learning and the relevance vector machine. *Journal of machine learning research*, *1*(Jun), 211–244.

Todd, P. M., & Dieckmann, A. (2005). Heuristics for ordering cue search in decision making, In *Advances in neural information processing systems*.

Todd, P. M., & Gigerenzer, G. (2000). Précis of simple heuristics that make us smart. *Behavioral and brain sciences*, *23*(5), 727–741.

Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current directions in psychological science*, *16*(3), 167–171.

Todd, P. M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. Oxford University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124–1131.

Van Ravenzwaaij, D., Moore, C. P., Lee, M. D., & Newell, B. R. (2014). A hierarchical bayesian modeling approach to searching and stopping in multi-attribute judgment. *Cognitive Science*, *38*(7), 1384–1405.

Van Rooij, I., Wright, C. D., & Wareham, T. (2012). Intractability and the use of heuristics in psychological explanations. *Synthese*, *187*(2), 471–487.

Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., Blundell, C., Kumaran, D., & Botvinick, M. (2016). Learning to reinforcement learn. *arXiv preprint arXiv:1611.05763*.

Zaslavsky, N., Kemp, C., Regier, T., & Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, *115*(31), 7937–7942.

Appendix A

Variational Inference Details

We update posterior distributions over weights after each observation using variational inference. The true posterior is approximated with a normal distribution $q(\mathbf{w}; \boldsymbol{\lambda}_t) = \mathcal{N}(\mathbf{w}; \boldsymbol{\mu}_t, \boldsymbol{\Psi}_t)$ and its parameters $\boldsymbol{\lambda}_t = (\boldsymbol{\mu}_t, \boldsymbol{\Psi}_t)$ are obtained through maximizing the evidence lower bound:

$$\mathcal{L}(\boldsymbol{\lambda}_t) = \mathbb{E}_{\mathbf{w} \sim q(\mathbf{w}; \boldsymbol{\lambda}_t)} \left[\log p(C_t = c_t | \mathbf{x}_t, \mathbf{w}) \right] - \mathrm{KL} \left[q(\mathbf{w}; \boldsymbol{\lambda}_t) || q(\mathbf{w}; \boldsymbol{\lambda}_{t-1}) \right] \tag{5}$$

The initial prior is set to a standard normal distribution $q(\mathbf{w}; \boldsymbol{\lambda}_0) = \mathcal{N}(0, \mathbf{I})$ and we furthermore adopt a mean field approximation, where posterior covariance matrices $\boldsymbol{\Psi}_t$ are restricted to be diagonal. To ensure positive semi-definite covariance matrices we parametrize them with logarithms of their standard deviations.

Equation 5 is maximized through gradient-based optimization using AmsGrad (Reddi et al., 2019) with a learning rate of 0.1. Training is stopped once the evidence lower bound function does not increase anymore over 10 steps or after 1000 total gradient steps.

The Kullback-Leibler divergence can be evaluated in closed-form assuming normal prior and posterior distributions. The expected log-likelihood term is approximated through 100 samples and we employ the reparametrization trick (Kingma & Welling, 2013) to obtain gradients with respect to the variational parameters $\boldsymbol{\lambda}_t$.

Appendix B

Meta-Learning Details

The architecture of our BMLI models consists of a Gated Recurrent Unit (GRU, Cho et al., 2014) with a hidden size of 128 units, followed by two linear transformations projecting to $\boldsymbol{\mu}_t \in \mathbb{R}^D$ and $\log \boldsymbol{\sigma}_t \in \mathbb{R}^D$ respectively. The latter are used to construct diagonal posterior covariance matrices $\boldsymbol{\Psi}_t$ as in the ideal observer model. Thus, $\boldsymbol{\lambda}_t = (\boldsymbol{\mu}_t, \boldsymbol{\Psi}_t)$ is a function of $(\mathbf{x}_1, c_1, \dots, \mathbf{x}_t, c_t)$. BMLI is optimized using the AmsGrad optimizer (Reddi et al., 2019) to minimize Equation 7:

$$\mathcal{L}(\boldsymbol{\Lambda}) = -\mathbb{E}_{q(\boldsymbol{\Theta};\boldsymbol{\Lambda})} \left[\sum_{t=1}^{T} \log p(C_t = c_t | \mathbf{x}_t, \boldsymbol{\lambda}_{t-1}, \boldsymbol{\Theta}) \right] + \beta \mathrm{KL} \left[q(\boldsymbol{\Theta}; \boldsymbol{\Lambda}) || p(\boldsymbol{\Theta}) \right] \tag{7}$$

Learning rates are set to $3 \cdot 10^{-4}$ and we train for 10^6 gradient steps; at this point the loss function has converged. Each model is initialized from a pretrained version without resource limitations and we increase β linearly over the first half of the training to the desired value.

The prior over network parameters corresponds to a variational dropout prior (Kingma et al., 2015; Molchanov et al., 2017). Like Molchanov et al. (2017) we use a normal distribution with diagonal covariance matrices as encoding distribution.

During training the expectation of the log-likelihood term is approximated through one sample from the encoding distribution $q(\boldsymbol{\Theta}; \boldsymbol{\Lambda})$ and we obtain gradients with respect to $\boldsymbol{\Lambda}$ using the reparametrization trick (Kingma & Welling, 2013). For evaluation we use an approximation based on 100 samples from the encoding distribution.

Appendix C

Power Analysis

Environments with continuous features can facilitate statistical analysis as fewer trials are needed to observe expected effects. To verify this hypothesis, we conducted a power analysis for an environment with continuous features and one for an environment, where features are dichotomized based on their median. Here, we present results from environments with known feature rankings and $T = 10$ decisions per task.

In both settings we compute how many tasks are on average required to distinguish the single cue heuristic from the ideal observer model, assuming that decisions are made by the single cue heuristic. In dichotomized environments ties between features of two options are likely, and hence we modify the single cue heuristic to make decisions based on the first feature that discriminates between both options.

We assume that decisions are made by the single cue heuristic and measure the support for the single cue heuristic over the ideal observer model on a specific task $\mathcal{D}$ by computing log-Bayes Factors (Kass & Raftery, 1995) between both strategies:

$$B(\mathcal{D}) = \sum_{t=1}^{T} \left[\int p(C_t = c_t | \mathbf{x}_t, w, m = \text{SC}) \log \left(\frac{p(C_t = c_t | \mathbf{x}_t, w, m = \text{SC})}{p(C_t = c_t | \mathbf{x}_t, \mathbf{w}, m = \text{IO})} \right) dc_t \right] \tag{9}$$

$$= \sum_{t=1}^{T} \text{KL} \left[p(C_t = 1 | \mathbf{x}_t, w, m = \text{SC}) || p(C_t = 1 | \mathbf{x}_t, \mathbf{w}, m = \text{IO}) \right] \tag{10}$$

and approximate the expectation over all possible tasks using $N = 10^5$ samples:

$$B = \mathbb{E}_{\mathcal{D}} \left[B(\mathcal{D}) \right] \approx \frac{1}{N} \sum_{\mathcal{D}} B(\mathcal{D}) \tag{11}$$

Because tasks are sampled independently from each other, we can multiply B by the total number of encountered tasks K to get expected log-Bayes Factors for an experiment with K tasks. Figure C1 shows this analysis for both continuous and dichotomized environments. We observe that it requires over four times more tasks to distinguish the single cue heuristic from an ideal observer model in environments with dichotomized features compared to one with continuous features.

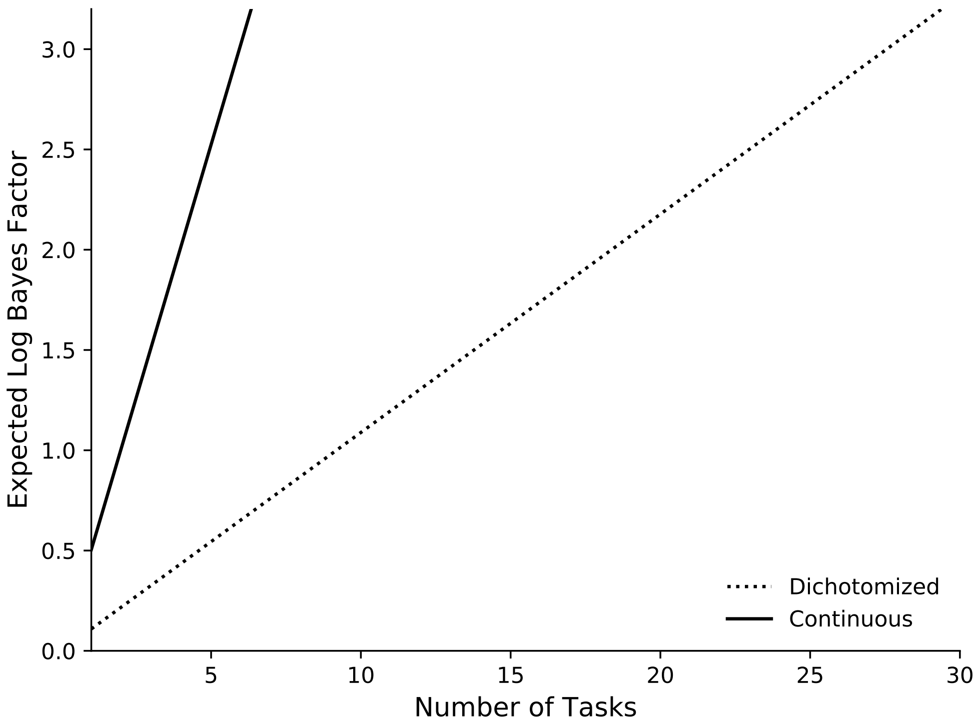

Figure C1

Power analysis for environments with known feature ranking. The plot illustrates how many tasks are on average required to distinguish the ideal observer model from the single cue heuristic, assuming that decisions are made by the single cue heuristic. We show results for both dichotomized environments (dotted) and environments with continuous features (solid).

Appendix D

Bayesian Model Comparison

We rely on Bayesian model comparisons (Bishop, 2006) to test which hypothesis accounts best for human decision making. Bayesian model comparison provides us with a principled tool for comparing evidence for different models. For the most part we perform separate comparisons for each participant in order to detected potential individual differences.

Let $\mathbf{X}^{(i)} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{KT}\}$ denote the set of all observed features and $\hat{\mathbf{c}}^{(i)} = \{\hat{c}_1, \ldots, \hat{c}_{KT}\}$ the set of corresponding decisions from a single participant i, and let $\mathbf{X}$ and $\hat{\mathbf{c}}$ denote the joint data for all participants. K corresponds to the total number of tasks and T to the number of decisions per task. Note, that we use $\hat{c}$ to refer to decisions made by participants and c to refer to ground truth labels. Using this data, we can compute the probability that a participant used strategy m through Bayes' rule:

$$p(m|\hat{\mathbf{c}}^{(i)}, \mathbf{X}^{(i)}) = \frac{p(\hat{\mathbf{c}}^{(i)}|\mathbf{X}^{(i)}, m)p(m)}{p(\hat{\mathbf{c}}^{(i)}|\mathbf{X}^{(i)})} \tag{12}$$

In our analysis we assume a uniform prior over all hypothesis and that the evidence factorizes over decisions. Model parameters are either set to maximize the evidence lower bound (Equation 5) on past observations or obtained from BMLI. In both cases they are not free parameters subject to the comparison.

For some analyses we additionally want to compare BMLI against other strategies. In such cases we additionally determine the level of resource limitations that best describes each participant. To account for fitting the additional parameter we use the Bayesian information criterion (Schwarz et al., 1978):

$$\log p(\hat{\mathbf{c}}^{(i)}|\mathbf{X}^{(i)}, m = \text{BMLI}) \approx -\frac{1}{2}\log KT + \max_{\beta} \log p(\hat{\mathbf{c}}^{(i)}, |\mathbf{X}^{(i)}, m = \text{BMLI}_{\beta}) + \text{const.} \tag{13}$$

where BMLI_{β} represents BMLI with a specific value of β and BMLI refers to the hypothesis that any resource rational inference algorithm was used.