

Towards a theory of heuristic and optimal planning for sequential information search

Jonathan D. Nelson<sup>1,2</sup>, Björn Meder<sup>1,3</sup>, Matt Jones<sup>4</sup>

<sup>1</sup> MPRG iSearch, Max Planck Institute for Human Development, Berlin, Germany

<sup>2</sup> University of Surrey, Guildford, UK

<sup>3</sup> Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin,  
Germany

<sup>4</sup> University of Colorado Boulder, USA

The authors contributed equally and share first authorship. Correspondence can be addressed to Jonathan Nelson (jonathan.d.nelson@gmail.com, +44 01483-68-890), Björn Meder (meder@mpib-berlin.mpg.de, +49 030-82406-239), or Matt Jones (mcj@colorado.edu, +1 303-735-1351).

### Abstract

How should tests (or queries, questions, or experiments) be selected? Does it matter if only a single test is allowed, or if a sequential test strategy can be planned in advance? This article contributes two sets of theoretical results bearing on these questions. First, for selecting a single test, several Optimal Experimental Design (OED) ideas have been proposed in statistics and other disciplines. The OED models are mathematically nontrivial. How is it that they often predict human behavior well? One possibility is that simple heuristics can approximate or exactly implement OED models. We prove that heuristics can identify the highest information value queries (as quantified by OED models) in several situations, thus providing a possible algorithmic-level theory of human behavior. Second, we address whether OED models are optimal for sequential search, as is frequently presumed. We consider the Person Game, a 20-questions scenario, as well as a two-category, binary feature scenario, both of which have been widely used in psychological research. In each task, we demonstrate via specific examples and extended computational simulations that neither the OED models nor the heuristics considered in the literature are optimal. Little research addresses human behavior in such situations. We call for experimental research into how people approach the sequential planning of tests, and theoretical research on what sequential planning procedures are most successful, and we offer a number of testable predictions for discriminating among candidate models.

*Keywords:* Sequential information search, optimal experimental design (OED) models, heuristics and optimality, split-half heuristic, maximum-entropy question heuristic, stepwise and multi-step question strategies

## 1. Introduction

A doctor cannot conduct every available medical test when diagnosing a patient. A scientist cannot conduct all potentially relevant experiments. A child cannot point at every object to inquire whether it is consistent with the meaning of a new word. But carefully chosen queries can greatly assist learning and decision making in all of these situations. (We use the term query in a general sense: it could be a scientific experiment, a question, an eye movement, etc.) Given that queries are costly (e.g., in terms of money, time, or potential harm to a patient), an important issue is which strategy for selecting queries is most efficient, to obtain a maximum amount of information with as few tests as possible.

Can the informativeness of different queries be precisely quantified? Some ideas for addressing this question began to develop in statistics in the 1950s (Good, 1950; Lindley, 1956); the general approach falls within Savage's (1954) Bayesian decision-theoretic framework. Several so-called *Optimal Experimental Design* (OED) models have been proposed to quantify the informational value of queries in this framework. OED models measure things like how much uncertainty has been reduced (information gain; Lindley, 1956), classification accuracy has been improved (probability gain; Baron, 1985), or beliefs have changed (e.g., impact; Wells & Lindsay, 1980). Related terms include *epistemic utility* and *quasi utility*, *sampling norm*, *value-of-information model*, and *Optimal Data Selection*.

How do people select queries? There has recently been an explosion of work in this area, including in the study of causal learning, development, machine learning, eye movements in visual perception, and medical decision making (reviewed by Coenen, Nelson & Gureckis, 2018). Although OED models were initially designed as normative benchmarks in statistics, and their calculation is mathematically nontrivial, they predict human information acquisition well in a variety of tasks (Nelson, 2005a). However, there are two reasons to suspect this is not the full story, and these are the two foci of the present paper.

First, it is not clear how people select queries that align with the predictions of particular OED models. One possibility is that the brain, without the person's awareness, does implement sophisticated query-selection strategies. This might be the most plausible for tasks such as selecting eye movements, which rely on evolutionarily old oculomotor neural circuits. But it seems less plausible in higher-level cognitive tasks. Consider the 20 questions game, an important task in cognitive and developmental research, in which the goal is to identify an unknown target item by asking yes-no questions about its features. This task presumably draws heavily on explicit reasoning, and thus limitations in working memory (and other cognitive constraints) make it unlikely that people are explicitly computing entropy and related quantities used by OED models. In this case, a promising avenue for reconciling statistical OED theory and psychology lies in exploring whether heuristic algorithms exist that people could use to make smart test-selection choices. In fact, particular heuristics for information acquisition have been shown to approximate (Klayman & Ha, 1987; Markant, Settles, & Gureckis, 2015; Slowiaczek, Klayman, Sherman, & Skov, 1992) or even exactly implement (Navarro & Perfors, 2011; Nelson, 2005a, 2005b) specific OED models. We extend these results by providing several new mathematical proofs on the relationships between heuristics and OED models.

A second issue pertaining to most extant OED models in psychology, and the second focus of this paper, is that these models are typically implemented in a *stepwise-optimal* fashion, meaning that they aim to achieve a maximum amount of information in the current time step. Such stepwise methods (also known as *greedy* or *myopic*) strongly reduce the computational complexity of the models, but because they do not plan ahead, they can be distinctly suboptimal in sequential search scenarios in which more than one query can be conducted in succession. The psychological literature on information search typically treats the criterion of maximizing stepwise information gain (which we formally define later on) as a normative benchmark for human information acquisition (e.g., Navarro & Perfors, 2011;

Ruggeri, Lombrozo, Griffiths & Xu, 2015), but seldom addresses whether in fact it generally leads to maximum efficiency in sequential search. The key problem, as we see it, is that researchers have either failed to consider multistep settings, or they have implicitly assumed such settings are governed by the same normative principles as one-shot tasks. Also contributing to the confusion is inadequate characterization of the particular task environment that is considered, in particular whether questions pertaining to arbitrary subsets are allowed, and whether hypotheses are equally likely a priori or not. For many tasks of interest, the “optimal” query as quantified with a stepwise-OED method is not necessarily the best first query if more than one piece of information can be acquired. Thus, prior confusion in the literature notwithstanding, a critical theoretical and empirical issue is when stepwise methods fail, and what *multistep* procedures make it possible to plan more than one query at a time in a way that yields an efficient sequence of questions.

## 2. Goals and Scope

What is known about people's ability to identify efficient sequential question strategies, in circumstances where the best first query in a sequence would not be the most useful standalone query? Very little is known about this. We suggest that this blind spot in the literature exists because the long-run suboptimality of stepwise-optimal methods is not widely appreciated in the hypothesis-testing literature in which OED models are used. This situation with research on pure information-gathering contrasts with research in reward-based tasks, where the tension between immediate outcomes and multistep considerations has long been appreciated (Bellman, 1957; Newell & Simon, 1959). We thus suggest that determining which stepwise method is used (a highly studied issue) may be much less important than identifying the circumstances in which people plan ahead in sequential information acquisition, the algorithms they use to plan, and the time horizon of their planning. Although the present paper is theoretically focused, for each problem considered we highlight how our

theoretical results open up new important areas for future research on the psychology of human information acquisition, including specific predictions that differentiate between potential models.

The rest of this paper is structured as follows. First, we briefly introduce the theoretical foundations of information search from the perspective of inductive inference, focusing on key OED models and how they have been used to investigate human information acquisition. Second, we show that simple heuristics can correspond exactly to particular OED models on two prominent tasks from research on human information acquisition. However, we show in the subsequent sections that neither stepwise OED models nor simple heuristics can be relied on to identify the most efficient strategy in sequential search contexts.

The first task we consider is the *Person Game*, a variant of the 20 questions game in which the goal is to identify an unknown target hypothesis with as few binary and deterministic (yes-no) questions as possible. This task was a focus of early developmental psychological research on human information acquisition (e.g., Mosher & Hornsby, 1966) and is an active area of current empirical research. It is also an important problem in applied and theoretical computer science. For instance, Hyafil and Rivest (1976) proved that the problem of finding the most efficient search strategy in this task is NP-complete. Our contributions include (i) new proofs on the relationship of heuristics and stepwise OED models; (ii) precise mathematical characterization of the implications of different question types central to the developmental literature; (iii) demonstration that stepwise-optimal procedures can be distinctly suboptimal with respect to long-run efficiency; (iv) new results on stepwise and multistep probability gain (accuracy maximization) on generalized versions of the Person Game; and (v) characterization of possible heuristic procedures for calculating multistep probability gain. We further show via simulations (vi) that the suboptimality of stepwise procedures occurs in many different kinds of Person Game environments.

We then consider a second task, the *Planet Vuma* scenario, which has been used in several social and cognitive psychology experiments on human information acquisition (e.g., Skov & Sherman, 1986). This task is defined by two categories and two or more probabilistic binary features; searchers' job is to identify which feature is most informative for classification. This task has been used to differentiate which of several OED models best predicts human information acquisition behavior (Nelson, 2005a). Results suggested that human behavior is best described by maximizing probability gain, at least when only a single query can be conducted (Nelson et al., 2010). The psychological centrality of probability gain is consistent with learning models' focus on error reduction for driving attention (Kruschke, 2001). We provide new results showing (i) that if two queries can be made, rather than a single query as in past research, which query is the best first query for maximizing accuracy (i.e., maximizing probability gain) can actually switch, and (ii) that this divergence occurs in a wide range of circumstances.

The Person Game and Planet Vuma are complex in different ways: the Person Game includes many hypotheses; Planet Vuma includes nondeterministic feature likelihoods, such that figuring out the true hypothesis with certainty may not be possible. Real world tasks tend to include both of these kinds of complexities. For instance, medical diagnosis is characterized by many possible diseases (hypotheses) and nondeterministic tests (with some false positive and false negative results). The fact that stepwise methods fail in both the Person Game and in Planet Vuma suggests that stepwise methods may be suboptimal in a wide variety of tasks important to the psychological literature. We conclude by discussing how ideas from computer science, statistics, and psychology are critical for thinking about theory and experiments to best characterize and investigate human sequential search.

### **3. Human information acquisition: History and heuristics**

Early work on human information search took a logico-deductive perspective, inspired by Popper's (1959) falsificationist philosophy of science. A prominent example is Wason's (1966, 1968) card selection task, in which participants are asked to find out whether a conditional rule holds; and Wason's (1960) rule-learning (2-4-6) task, in which the goal is to identify an unknown rule by generating triplets of numbers and inquiring whether they are consistent with the rule. Wason and others concluded (before and then in parallel with the heuristics-and-biases tradition in psychology; Tversky & Kahneman, 1974; Gilovich & Griffin, 2002) that human intuitions about question selection were fraught with suboptimalities, such as various forms of "confirmation bias" (Klayman, 1995; Klayman & Ha, 1987; Nickerson, 1996).

In the 1980s, psychologists started using OED principles as normative models for question selection (Baron, 1985; Skov & Sherman, 1986; Slowiaczek, Klayman, Sherman & Skov, 1992; Trope & Bassok, 1982). Baron, Beattie and Hershey (1988) analyzed information acquisition in a medical diagnosis task from this perspective. Oaksford and Chater (1994, 1996, 2003) analyzed the Wason selection task from the perspective of inductive inference, positing that subjects have the goal of finding out whether it is more likely that a rule holds or not, rather than trying to falsify the rule. Oaksford and Chater used expected reduction in Shannon (1948) entropy to quantify cards' information value. Nelson (2005a) showed that similar results are obtained if probability gain or impact, instead of information gain, is used.

### **4. Optimal Experimental Design (OED) models**

Each OED model quantifies the expected usefulness of a possible query for discriminating among the hypotheses under consideration. The OED models are defined only in terms of the relevant probability model, i.e., the set of priors and likelihoods for the task at hand. Thus, the OED approach is appropriate for situations in which the goal is pure



knowledge, or in which the actual external payoffs unknown. However, the mathematical framework can be extended to incorporate situation-specific utilities, such as the costs and benefits of alternative treatment decisions in medical scenarios, for quantifying queries' usefulness (Baron & Hershey, 1988; Markant & Guereckis, 2012; Meder & Nelson, 2012; Pauker & Kassirer, 1980; Raiffa & Schlaifer, 1961.)

We use the following notation:  $Q$  denotes a query, which can yield a finite number  $s$  of possible outcomes (answers),  $q_1, \dots, q_s$ . In a categorization task, a query  $Q$  could be to view a physical feature of an object, with the various  $q_j$  denoting the possible values of that feature. In a medical diagnosis task,  $Q$  could be a medical test, with the various  $q_j$  denoting the possible test outcomes (e.g., a positive or negative result).  $H = \{h_1, \dots, h_n\}$  refers to the hypothesis space under consideration, with  $h_1, \dots, h_n$  denoting the  $n$  individual hypotheses, e.g., the  $n$  possible object categories or disease diagnoses.  $P(h_i)$  is the prior probability of a hypothesis;  $P(q_j | h_i)$  is the *likelihood* of a datum given a hypothesis (e.g., the true or false positive rate of a medical test). The marginal probability of each datum is given by  $P(q_j) = \sum_i P(h_i) \cdot P(q_j | h_i)$ . Finally,  $P(h_i | q_j)$  is the *posterior probability* of a particular hypothesis given a particular piece of evidence (e.g., the probability that the patient has a particular disease given a positive test result), which can be calculated using Bayes' (1763 / 1958) theorem.

A query  $Q$ 's expected usefulness,  $eu(Q)$ , is computed by averaging the usefulness of each of its possible outcomes,  $u(q_j)$ , weighted according to each outcome's probability of occurrence:

$$eu(Q) = \sum_{j=1}^s P(q_j) u(q_j). \quad (1)$$

The key distinction between different OED models is how the usefulness of a particular outcome,  $u(q_j)$ , is quantified. Various models from statistics, philosophy of science, medical decision making, and psychology offer different solutions to this question (Benish, 1999; Crupi & Tentori, 2014; Crupi, Nelson, Meder, Cevolani, & Tentori, 2018; Good, 1950; Lindley, 1956; Oaksford & Chater, 1994, 1996, 2003). Importantly, the various OED models are *not* equivalent (Nelson, 2005a), and they rank the usefulness of available questions differently in important cases. We here focus on three prominent models, namely impact, information gain, and probability gain, because of their centrality in psychological literature.

#### 4.1 Probability gain

*Probability gain* (PG; Baron, 1985) is defined as the improvement in the probability of making a correct guess, assuming that one always guesses the most probable hypothesis given all knowledge to date. An answer's probability gain, which can be positive or negative, is the accuracy after asking the question minus the accuracy that could have been obtained without asking the question:<sup>1</sup>

$$u_{PG}(q_j) = \max_{1 \leq i \leq n} P(h_i | q_j) - \max_{1 \leq i \leq n} P(h_i). \quad (2)$$

Probability gain maximization successfully predicts subjects' questions on experience-based probabilistic categorization tasks (Nelson et al., 2010) and has been used as a normative benchmark in medical diagnosis test selection (Baron, Beattie & Hershey, 1988); to predict human eye movements on visual search tasks (Najemnik & Geisler, 2005; but see Morvan & Maloney, 2012); to describe a learner's goals in a Bayesian formulation of active associative

---

<sup>1</sup> We follow standard order of operations, in which Sigma summation symbols extend up to, but do not include, the next plus or minus symbol. We give max and min operators the same precedence level as Sigma summation symbols. Thus in Equation 2,  $\max P(h_i | q_j) - \max P(h_i)$  is read as  $(\max P(h_i | q_j)) - (\max P(h_i))$ . When indexing max or min operators, the shorthand  $1 \leq i \leq n$  means  $i \in \{1, 2, \dots, n\}$ .

learning (Kruschke, 2008); and as a rational benchmark in a Bayesian analysis of the pseudodiagnosticity paradigm (Crupi, Tentori, & Lombardi, 2009).

## 4.2 Impact

*Impact* (Imp; Wells & Lindsay, 1980; Nelson, 2005a) uses the absolute change (i.e., difference) between prior and posterior probabilities as the measure of how much information is provided about each hypothesis. The impact of an answer is thus given by

$$u_{\text{Imp}}(q_j) = \sum_{i=1}^n |P(h_i|q_j) - P(h_i)|. \quad (3)$$

Nelson (2005a) proved in the case of binary hypotheses and queries that a simple strategy, the *likelihood difference heuristic* (described below in Section 9), can be used to identify the highest-impact query, demonstrating how simple heuristics can exactly implement OED models. The likelihood difference heuristic has been reported to have been used by some subjects in two-category, binary-feature tasks (Planet Vuma and variant tasks; e.g., Skov & Sherman, 1986; Nelson, 2005a). People seem most apt to use the likelihood difference heuristic when the relevant probabilities are presented using the words-and-numeric *standard probability format* (as defined by Gigerenzer & Hoffrage, 1995).

## 4.3 Information gain

Information gain (IG; Lindley, 1956) quantifies the value of an answer as the reduction in uncertainty over the hypotheses under consideration, where uncertainty is measured with Shannon (1948) entropy. Accordingly, the usefulness of a datum is the entropy before asking a question, minus the entropy after obtaining the answer:

$$u_{\text{IG}}(q_j) = \text{ent}(H) - \text{ent}(H|q_j) = \sum_{i=1}^n P(h_i) \log_2 \frac{1}{P(h_i)} - \sum_{i=1}^n P(h_i|q_j) \log_2 \frac{1}{P(h_i|q_j)} \quad (5)$$

where  $\text{ent}(H)$  denotes the (Shannon) entropy of the prior distribution and  $\text{ent}(H | q_j)$  denotes the entropy of the posterior distribution given outcome  $q_j$ . (Note that the base of the logarithm is arbitrary; we here use  $\log_2$  such that the unit is the bit.)

Information gain has been used to predict human question selection (e.g., Austerweil & Griffiths, 2011; Bramley, Lagnado & Speekenbrink, 2015; Nelson, Divjak, Gudmundsdottir, Martignon & Meder, 2014; Nelson, Tenenbaum & Movellan, 2001; Oaksford & Chater, 1994, 1996, 2003; Steyvers, Tenenbaum, Wagenmakers & Blum, 2003), to guide a robot's eye movements (Denzler & Brown, 2002), to predict human eye movements (Legge, Klitz, & Tjan, 1997; Najemnik & Geisler, 2005; Renninger, Coughlan, Verghese & Malik, 2005), and to predict firing of individual neurons (Bromberg-Martin & Hikosaka, 2009; Nakamura, 2006). Information gain and related measures have also been used in machine learning approaches to classification tree induction (Breiman, Friedman, & Olshen, 1984; Quinlan, 1986). The *Adaptive Design Optimization* (ADO) approach to the design of psychological experiments (Myung, Cavagnaro, & Pitt, 2013) is also based on optimizing the information gain of each manipulation (mini experiment), in the next time step.

#### 4.4 The maximum entropy question heuristic and information gain

The maximum entropy question heuristic (also called *uncertainty sampling*) is the strategy of asking the question with the highest question entropy,  $\text{ent}(Q) = -\sum_{j=1}^s P(q_j) \log_2 P(q_j)$ . There is a close relationship between the uncertainty in a question's outcome and that question's information value with respect to the hypotheses of interest. This connection has been explored previously in machine learning (Garey & Graham, 1972) and more recently in psychology (Nelson, 2005b; Markant, Settles, & Gureckis, 2015). For completeness we restate the formal relationship, using the terminology of the present manuscript:

**Theorem 1:** *If likelihoods are deterministic, then a query's entropy is the same as its expected information gain; i.e., if  $\text{ent}(Q|H) = 0$ , then  $eu_{IG}(Q) = \text{ent}(Q)$ .* The proof is given in Appendix A.

Theorem 1 states that the maximum entropy query corresponds to the maximum information gain query in every domain in which each hypothesis makes deterministic predictions for the outcomes of all queries. Intuitively, the condition of deterministic likelihoods, i.e.  $\text{ent}(Q|H) = 0$ , is equivalent to the condition that all uncertainty in  $Q$  is due to uncertainty in  $H$  (rather than to noise or other extraneous factors), and therefore all information carried by  $Q$  is information that is relevant to inferring  $H$ . This relationship holds if priors are uniform or not, and irrespective of whether questions are binary or many-valued.

Examples of tasks covered by Theorem 1 include Wason's (1960) rule learning task, Klayman and Ha's (1989) number and geography concept tasks, 20 questions-type games (Kachergis, Rhodes & Gureckis, 2016; Nelson et al., 2014; Ruggeri & Lombrozo, 2015), and Markant, Settles and Gureckis's (2015) three-category-learning task. On these tasks, a learner who wishes to maximize stepwise information gain only has to identify which query's outcomes he or she is maximally uncertain about, given his or her current beliefs. For instance, Nelson, Tenenbaum and Movellan's (2001) active number concept learning task includes tens of thousands of hypotheses and unequal priors, but only 100 binary queries. If a learner can identify the maximum entropy question among the 100 available queries, he or she will necessarily select the highest information gain question. Tasks that do not fall under the umbrella of Theorem 1 include Meier and Blair's (2013) scenario, one of the scenarios considered by Navarro and Perfors (2011), and the Planet Vuma scenario (Skov & Sherman, 1986). Navarro and Perfors' results also demonstrate that there are cases in which uncertainty sampling maximizes information gain, even if the hypotheses' predictions are not deterministic.

## 5. From one-shot search decisions to sequential search

In psychology, the various OED models have mostly been used to determine only the single next query, even in tasks in which multiple queries can be conducted before a decision must be made. How, from a normative standpoint, should a sequence of queries be selected? Is it adequate to use OED principles in a stepwise way, namely to consider just the next time step when deciding which query to conduct? The answer to the latter question is negative: Identifying the maximally efficient series of queries in general requires consideration of all future time steps at once. Planning a sequence of tests is interesting mathematically as well as psychologically, because the number of possible question trees grows super-exponentially with the number of available questions (see Appendix B), rendering the problem of finding the globally optimal (most efficient) tree NP-complete (Hyafil & Rivest, 1976). The computer science literature includes numerous articles on sequential search tasks (e.g., Berghman, Goossens, & Leus, 2009; Garey, 1972; Geman & Jedynek, 2001; Golovin & Krause, 2011; Hyafil & Rivest, 1976).

Some scenarios in the psychological literature on information search have demonstrated from a theoretical standpoint the suboptimality of stepwise methods. These tasks have typically been fairly complex, with nondeterministic likelihoods and dozens of possible hypotheses (Bramley et al., 2015; Najemnik & Geisler, 2005). Moreover, these tasks were not designed to establish whether people can plan sequential information search optimally. In other domains, the importance of sequential planning is well appreciated, including problem solving (Newell & Simon, 1959), sequential decision making (e.g., Hotaling & Busemeyer, 2012; Gureckis & Love, 2009; Naw, Niv, & Dayan, 2005), and bandit tasks with exploration-exploitation tradeoffs (e.g., Navarro, Newell & Schulze, 2016; Wu, Schulz, Speekenbrink, Nelson & Meder, 2018; Zhang & Yu, 2013). By contrast, papers on pure information-gathering tasks, without external payoffs (e.g., Navarro & Perfors, 2011; Ruggeri &

Lombrozo, 2015), have tended to presume that stepwise methods are also optimal in sequential search scenarios.

We demonstrate that stepwise OED methods can be suboptimal in two paradigmatic human information acquisition tasks. Our analyses provide a foundation for future empirical research on human sequential planning in information search tasks. We highlight a number of specific issues that are ripe for experimentation as we go along.

## 6. The Person Game

The *Person Game* is a 20 questions-type sequential search task in which the goal is to identify an unknown target object from among a specified set, asking as few yes-no questions about its features as possible. Variants of this paradigm were used to investigate children and adults' information search behavior in the 1970s and 1980s (e.g., Denney & Denney, 1973; Eimas 1970; Siegler, 1977; Thornton, 1982). These tasks are undergoing a renaissance, as recent papers have connected them to stepwise-OED models of the value of information, especially stepwise information gain (Kachergis, Rhodes & Gureckis, 2016; Meder, Nelson, Jones & Ruggeri, 2018; Navarro & Perfors, 2011; Nelson, Divjak, Gudmundsdottir, Martignon & Meder, 2014; Ruggeri & Lombrozo, 2015; Ruggeri, Lombrozo, Griffiths, & Xu, 2016). Importantly, the Person Game task also offers rich opportunities to explore what is optimal for sequential search.

The Person Game, as we define it, is a non-strategic, single-player task. In the Person Game, there are  $n$  items (persons: hypotheses  $h_1$  through  $h_n$ ) and  $m$  binary features (queries  $Q_1$  through  $Q_m$ ), as in the example in Table 1 and Figure 1. (Novel monsters, rather than persons, are used as example stimuli in Figure 1, so that arbitrary relationships among features and items can be conveyed.) Each item is uniquely characterized by whether or not it has each of the  $m$  features (i.e., by a binary vector of length  $m$ ). A target item is selected at random with equal probability ( $1/n$  for each item), and the player's task is to identify the true hypothesis

with the minimum number of questions. Importantly, questions can only be about binary features from a prespecified list, each pertaining to whether some feature is present or absent in the target. It is not possible to ask questions pertaining to arbitrary subsets, such as “Is it Maria, Frank, or Ana?”. This limitation on the set of available queries corresponds to the characteristics of real-world situations (for example in medicine or science) where there is not always an available query with exactly the desired characteristics. As will be seen below, it is because of this limitation that planning an efficient sequence of questions in this task is not reducible to stepwise methods. Our definition of the Person Game exactly corresponds to Hyafil and Rivest’s (1976) scenario, and thus their (and subsequent) computer science results are directly relevant. We further stipulate that the set of features includes one feature that is unique for each of the  $n$  items, ensuring that it is always possible to identify the true item after enough questions have been asked.

Early papers on the Person Game and related tasks categorized questions according to different question types: *constraint questions* and *hypothesis-scanning* (or *name*) *questions* (Mosher & Hornsby, 1966). Constraint questions split the items so that no matter which answer is obtained, at least two items will be eliminated. (Constraint questions are possible only if there are at least four persons. If there are three or fewer persons, then all non-redundant questions are equally useful.) Hypothesis-scanning questions, which we will also refer to as *name questions*, isolate a single item (e.g., “Is it Maria?”). Constraint questions are presumed in the literature to be superior to hypothesis-scanning questions. Whether this presumption is correct is an issue that we address later on. An additional type of question, the *pseudoconstraint* question (which queries a feature present or absent in just a single item), is mathematically equivalent to the name question, so we do not discuss it here.



## 6.1 Stepwise-OED Strategies, and the split-half heuristic, in the Person Game

Before addressing whether stepwise strategies are optimal, we first address the question of how people could do something like maximize stepwise information gain—which is based on differences of probability-weighted expectations of averages of logarithms—in the first place. Are there simple heuristic strategies that people could use to identify the questions prescribed by one or more OED models?

### 6.1.1 The split-half heuristic

Of special interest in psychological research on the Person Game and related tasks is the *split-half heuristic* (Navarro & Perfors, 2011; Nelson et al., 2014). If half of the persons in the current set have a hat, and all other features are less equally distributed, this strategy would ask about the hat feature. If no feature splits the items exactly in half, the heuristic chooses a feature that comes as close as possible to a 50:50 split. Suppose that  $Q_1$  splits the current items more evenly than does  $Q_2$ . We will say that  $Q_1$  is more *splithalfy* than  $Q_2$ . Thus, the split-half heuristic asks the question with the greatest *splithalfiness*. In the example Person Game environment shown in Table 1,  $Q_1$  is unique in giving exactly a 4:4 split and thus is the question that the split-half strategy asks first. We also refer to the splithalfiness of a question at a non-initial step of the game (i.e., after one or more questions have been asked), based on how it splits the remaining hypotheses. Many experimental papers have reported that people (especially adults; this behavior tends to increase with age) often behave in accordance with the split-half heuristic (e.g., Eimas, 1970; Nelson et al., 2014; Siegler, 1977; Thornton, 1982). In the next subsection, we provide new mathematical results on the relationship of splithalfiness to the usefulness of questions as evaluated by stepwise OED models.

### 6.1.2 The split-half heuristic selects the maximum impact question, and the maximum information gain question, in the Person Game

We provide a new result showing that the expected impact (absolute change in beliefs) of a question in the Person Game is monotonically related to how close the probability of a yes answer is to .5, irrespective of the total number of persons (hypotheses)  $n$ .

**Theorem 2:** *In the Person Game, a question's expected impact,  $eu_{Imp}(Q)$ , is equal to  $4P(q_{yes})P(q_{no}) = 4P(q_{yes})[1-P(q_{yes})]$ . Consequently, the split-half strategy invariably selects the highest-impact question. The proof is given in Appendix A.*

Note that the expected impact of a question does not depend on the number of items ( $n$ ) per se, but only on  $P(q_{yes})$ , where  $P(q_{yes}) = n_{yes}/n$  and  $P(q_{no}) = 1 - P(q_{yes})$ . Importantly, the expression for  $eu_{Imp}(Q)$  given in Theorem 2 is a two-degree polynomial function of  $P(q_{yes})$ . The function is continuously differentiable; it has its maximum at  $P(q_{yes}) = .5$  and decreases symmetrically and monotonically from that point in both directions (Figure 2). Thus, the split-half heuristic invariably selects the highest-impact question in the Person Game, even if the searcher does not make any explicit reference to the equations for calculating impact. This result implies that stepwise impact considers  $Q_1$  to be the most useful first query in the environment shown in Table 1, as  $Q_1$  is the only question that splits the items exactly 50:50. A similar result relates the split-half heuristic and stepwise information gain:

**Theorem 3:** *In the Person Game, the split-half heuristic maximizes stepwise expected information gain. This result was previously given by Navarro and Perfors (2011). We derive it more concisely, by application of Theorem 1, in Appendix A.*

The key point in the proof of Theorem 3 is that the split-half heuristic invariably queries the maximum entropy question: Likelihoods are deterministic in the Person Game, i.e., each person (hypothesis) possesses each feature with probability 1 or 0. Thus, Theorem 1 applies and the split-half heuristic identifies the maximum information gain question.

Theorems 2 and 3 provide a path for addressing the prima facie implausibility of the idea that people would explicitly carry out the calculations entailed by an OED model. Irrespective of whether a learner’s goal is to maximize stepwise impact or stepwise information gain, it is not necessary to explicitly calculate weighted averages of logarithms or anything else; at least in the Person Game, a simple heuristic does the job.

## 6.2 Stepwise probability gain is indifferent among questions in the Person Game

Probability gain quantifies the value of a query as a function of the improvement in classification accuracy. We show here that if questions are selected in a stepwise manner, then all questions have the same probability gain (provided they are non-redundant, i.e. that  $0 < P(q_{\text{yes}}) < 1$  and hence the answer will induce at least some change in our beliefs). Surprisingly, this invariance holds *irrespective* of how splithalfy the question is, for instance whether the question gives a 50:50 split, a 40:60 split, or even a highly unequal 1:99 split.

**Theorem 4:** *In the Person Game, where there are  $n$  items in the current set, all informative questions have expected probability gain  $1/n$ :*

$$eu_{PG}(Q) = \begin{cases} 1/n & 0 < P(q_{\text{yes}}) < 1 \\ 0 & \text{otherwise.} \end{cases}$$

The proof is given in Appendix A. *Every* informative (non-redundant) question has the same probability gain, specifically  $1/n$ , *irrespective* of the proportion of objects with the feature. For instance, in the environment in Table 1, probability gain considers the initial question “Is Feature  $Q_1$  present?” (4:4 split) equally useful as the initial question “Is Feature  $Q_7$  present?” (1:7 split), whereas both information gain and impact consider  $Q_1$  as most useful and  $Q_7$  as tied for least useful. Note also that this property holds at any point in the game (i.e., after any number of questions have been asked). This can easily be seen by defining  $n'$  as the number of hypotheses consistent with the answers received thus far, and  $n'_{\text{yes}}$  and  $n'_{\text{no}}$  as the numbers of these remaining persons (hypotheses) that do and do not have the feature. Mirroring the proof of Theorem 4, all informative questions will have probability gain of  $1/n'$ .

Does this result imply that probability gain (classification accuracy) is a “bad” goal to have, or a bad OED model? To the contrary: if your goal is classification accuracy and you can only ask one question before making a guess, Theorem 4 shows that you don’t need to look for a question with maximum splithaleness. As soon as you find a question whose answer is not already known, even if that question produces a lopsided 1:99 split, you should stop searching and ask that question, because all informative questions increase your expected probability of correctly guessing the target by exactly  $1/n$ .

Why is Theorem 4 counterintuitive? One possible explanation is that people’s intuitions about the value of information in this task are closer to information gain than to probability gain (notwithstanding Nelson et al.’s 2010 results, from a different task). Another possible explanation is that people’s intuitions are adapted somewhat to sequential-search scenarios, even in nominally one-shot tasks. As we show below, probability gain does differentiate among questions when it is adapted for sequential search, to take into account two or more steps in advance. These contrary possible explanations lead to interesting issues to explore in sequential-search tasks with human subjects, as we elaborate later on.

### **6.3 Stepwise methods can be suboptimal in the Person Game**

In light of the results above, showing that the split-half heuristic, maximal stepwise impact, and maximal stepwise information gain are all equivalent in the Person Game, we henceforth refer to these collectively as the *stepwise methods*. (Stepwise probability gain is indifferent among the questions, which is why we exclude it from our definition of the stepwise methods.) How do the stepwise methods behave when applied iteratively in a sequential task? Table 1 gives the stepwise-OED models’ values of each possible first question for the environment in Table 1, showing that stepwise impact and stepwise information gain agree with the split-half heuristic in preferring  $Q_1$ .

Is  $Q_1$  indeed the best first query? Here we address this issue, with respect to the goal of identifying the unknown target hypothesis with as few questions as possible, on average. (Expected value seems the simplest way to implement a preference for reaching certainty more quickly, but other goals, such as exponential temporal discounting, could also be considered.) To formalize this question, we first define the *expected path length* of a binary decision (question) tree. The expected path length is the number of queries the tree takes to identify the target item, averaged over all targets. Under this measure, the optimal strategy is to use a tree (not necessarily unique) that has the lowest expected path length, and the best first question is the first question (the root node) in that tree. The stepwise-optimal-strategy tree (Figure 3a, top) has an expected path length of 3.25 questions. The globally optimal (most efficient) tree (Figure 3b, top), however, starts with  $Q_2$  and needs only 3.125 questions on average. (As discussed below, two trees tie for globally optimal; the other starts with  $Q_3$ .)

This example shows that the presumption in the psychological information search literature that maximizing stepwise information gain minimizes the expected number of questions in the Person Game is incorrect. Information gain and other OED models were envisioned in statistics (e.g., Good, 1950; Lindley, 1956) as reasonable quasi-utility functions to use for selection of individual experiments. Since then, research in computer science has investigated how well these and other greedy algorithms perform in multistep tasks. Formal results based on the mathematical concept of submodular functions (Nemhauser, Wolsey, & Fisher, 1978) have proven lower bounds for the performance of greedy query-selection algorithms (Golovin & Krause, 2011). In the case of the Person Game, these results imply the expected path length for the stepwise methods can be no greater than  $\ln(n) + 1$  times the expected path length of the globally optimal tree. Similar results hold for tasks with nondeterministic likelihoods (Golovin, Krause, & Ray, 2010), although under different technical assumptions than are made here (e.g., in the Planet Vuma task below). However, a factor of  $\ln(n) + 1$  can be a large difference; even for relatively small environments it means

stepwise methods could require several times more queries than necessary. Moreover, it has also been proven that, for general search tasks with deterministic likelihoods, it is intractable (NP-hard) to produce decision trees with expected path length guaranteed to be less than  $\ln(n)$  times the optimal (Chakaravarthy, Pandit, Roy, Awasthi, & Mohania, 2007).

To understand why this result conflicts with intuition and with previous claims in the psychological literature, one must be careful about the precise nature of the task. In the Person Game, the list of available features (and hence allowable questions) is typically a small subset of the set of possible bipartite partitions of the items—in our example, 11 (the seven features shown in Table 1 plus the four other hypothesis-scanning questions) versus 127 (all nontrivial splittings of the 8 items). If arbitrary questions can be asked, or equivalently if there is a feature corresponding to every possible bipartite partition of the items, then the stepwise methods are globally optimal. In this case, the Huffman (1952) limit describes the expected number of questions required by the optimal strategy. We refer to this special case of the Person Game as the *Number Game* (Nelson et al., 2014). In the Number Game, the goal is to identify an unknown integer between, say, 1 and 16 (inclusive). Arbitrary questions are allowed, such as “Is the number equal to 2, 3, 5, 6, or 13?”, which enables a searcher to always partition the remaining hypotheses (i.e., numbers not yet eliminated) at each step into maximally equal subsets. Note that, as a special case of the Person Game, the Number Game also satisfies the conditions of Theorems 2, 3, and 4. Figure 4 illustrates the relationship between the Person Game, the Number Game, and other possible 20 questions-game tasks. Some of these relationships, for instance the implications if priors are unequal as in the Generalized Person Game, will be discussed later on.

In the Person Game, because it will not always be possible to ask a question that splits the remaining hypotheses close to 50:50, one needs to plan ahead and choose queries that not only will provide useful immediate information, but will also allow for useful queries at later steps. In our example (Table 1), question  $Q_1$  is maximally splithalfy, but suboptimal because

it leaves no highly splithalfy options for the second query, where the best one can do is a 3:1 split. By contrast, if starting with  $Q_2$  or  $Q_3$  and ending up in the bigger (more likely) set, one can still ask a 2:3 question.

### 6.3.1 Relation to Markov Decision Processes

Another perspective on sequential information search comes from the formalism of Markov decision processes (MDPs; Howard, 1960). MDPs are the basis for modern discrete-time control theory as well as computational models of reinforcement learning in psychology and machine learning (Sutton & Barto, 1998). An MDP is a formal task environment comprising a set of states, available actions within each state, transition probabilities for the next state given the current state and action, and a reward function that can depend on the current state, action, and next state. The agent's goal is to learn a policy, meaning a rule for selecting an action in each state, that maximizes reward accumulated across time steps.

Information search can be cast as a partially observable MDP (POMDP), an extension wherein the agent does not know the true state and must infer it from generally stochastic observations (Aström, 1965). Optimal solution of a POMDP involves maintaining a *belief state*, which is the posterior distribution over the identity of the true state conditioned on all past actions and observations (Kaelbling, Littman, & Cassandra, 1998). The belief state itself behaves according to a standard MDP, in that it contains all the information available to the agent for predicting the reward and the next belief state that would result from any candidate action. In a task of pure information search, one can treat the target as the true state (which is constant within each episode), the available queries as the actions the agent can perform, and the results of queries as the observations (Butko & Movellan, 2010). The belief state is the posterior over target hypotheses, updated after each new observation according to Bayes' rule. Note that the agent can determine the outcome probabilities for any candidate query (e.g., probabilities of positive and negative answers) by marginalizing over the belief state, and thus the transition probabilities for the subsequent belief state are known in advance. In other

words, the agent can plan ahead in the space of belief states to evaluate different possible question trees. This is true in nondeterministic search tasks, such as the Planet Vuma task considered later in this paper, as well as in deterministic tasks like the Person Game. In the remainder of this section, we use “state” to refer to belief state.

This picture is particularly simple in environments with deterministic likelihoods, such as the Person Game. There, a state can be identified with the set of remaining hypotheses or targets that are consistent with all answers received thus far (known as the version space in the machine learning literature; Mitchell, 1982). The transition probabilities following any candidate query are the probabilities of positive and negative answers ( $n_{\text{yes}}^t/n^t$  and  $n_{\text{no}}^t/n^t$ , the superscript indicating items remaining at the current step,  $t$ ). After a new query is answered, the state is updated by intersection, removing all items inconsistent with that answer. If we define reward as  $-1$  on every step, then the agent’s goal is to identify the target with the fewest number of queries. Other interesting reward functions are possible as well, as discussed later in this section.

The optimal policy in an MDP satisfies Bellman’s optimality equation (Bellman, 1957), which specifies a recursive relationship between the policy and the *value* it induces for each state. The value of a state is defined as the total future reward expected following that state, assuming the agent acts according to the policy in question. A policy is globally optimal if and only if it is optimal with respect to the values it induces; that is, the action it chooses in any state maximizes the expectation of the immediate reward plus the value of the subsequent state. In a search task with constant  $-1$  reward, the value of any state is the negative expected path length from that state. In tasks like the Person Game, where each episode is guaranteed to end in finite time, the optimal policy and state values—that is, the globally optimal tree and path lengths—can be determined by backward induction, as detailed in Appendix B. In brief, once values have been determined for all states with fewer than  $m$  items remaining (for any  $m$ ), the optimal query and value can be determined for all states with exactly  $m$  items. This



process shows how looking ahead matters: The value of a query depends on the values of states it might lead to, which in turn depend on the queries available in those states (and so on). It also shows why stepwise methods are adequate in the Number Game, where arbitrary subsets of items can be queried. A symmetry argument implies that the value of any state depends only on the number of items in that state, and therefore the value of a query depends only on how it splits the items currently remaining.

The MDP or POMDP perspective on multistep information search offers avenues for more sophisticated models and algorithms, by bridging to the computational reinforcement learning literature. For example, one can define reward functions to match the OED models considered here: Information gain corresponds to negative entropy, probability gain corresponds to the max operator (i.e., the greatest probability among all hypotheses), and impact corresponds to total absolute change in probabilities of the hypotheses. The stepwise OED models can be obtained by valuing only immediate reward, which in the MDP framework corresponds to setting the *temporal discount parameter* to zero. However, positive values of this parameter, corresponding to valuing later steps, can yield sophisticated multistep methods. For example, Butko and Movellan (2010) used this approach to train an algorithm to optimize long-run information gain. They tested it on the visual search task of Najemnik and Geisler (2005) and found that it yielded better performance than Najemnik and Geisler's stepwise probability gain model. Although Butko and Movellan's work was cast as machine learning, it might offer a bridge to development of new psychological theories.

### **6.3.2 Experimental questions: suboptimality of the stepwise methods**

Figure 1 illustrates how the statistical structure of Table 1 could be presented to experiment participants in a “Monster Game”. The task would be to identify a single monster, chosen at random with equal probability from among the 8 monsters in Figure 1.

The shape question ( $Q_1$  in Table 1) gives a 4:4 split; it is selected by the stepwise methods. However, the color question ( $Q_2$ ) and the eye question ( $Q_3$ ), despite giving less

splithalfy 3:5 splits, tie for optimal in terms of expected path length. An environment like this would be useful for exploring whether, and the circumstances under which, people can identify the global usefulness of questions, separate from their immediate splithalfiness. Specifically, when do people start with a globally optimal question ( $Q_2$  or  $Q_3$ ) as opposed to the stepwise-optimal question ( $Q_1$ )? Although this has not been extensively explored in the literature, a first set of experiments (Meder et al., 2018) suggests that people may have a hard time planning beyond the next timestep in standard Person Game experimental manipulations. Would people be more sensitive to long-run efficiency if they were explicitly encouraged to look ahead? This latter question could be tested by manipulations that encourage people to consider what subsequent questions will be available, given possible answers to candidate first questions. Assuming people focus on splithalfiness, they could be asked to rate the splithalfiness of the questions that would remain available, given the possible answers to their first question.

A further issue is whether recognizing the conflict between stepwise and global optimality depends on perceiving the dependencies among features. A feature can be highly splithalfy but poor as a first question if it is highly correlated with other available features (see Table 1:  $Q_1$  has strong correlations with both  $Q_2$  and  $Q_3$ ). Thus, the query leaves no good follow-up questions that could give substantial additional information. The dependence of people's planning ability upon their familiarity with the statistical structure among the features could be tested using preliminary tasks such as having subjects arrange cards on a table in a way that their spatial layout corresponds to the partitions in an available question tree.

#### **6.4 Performance of the stepwise methods in different environments**

We have shown for one specific example that the stepwise methods are suboptimal. Is our example pathological or otherwise unusual? Or do the stepwise methods fail across a wide

range of Person Game environments? What happens if the environments are more complex? In the following, we address these issues via analytic and simulation results.

#### 6.4.1 Analytic results

It can be shown analytically that in particular situations the stepwise methods invariably identify the best first question.

*Theorem 5: If there are fewer than eight items in a Person Game environment, then the stepwise methods invariably select a first question with minimal expected path length. The proof is in Appendix A.*

*Theorem 6: If a Person Game environment has only two constraint questions, then the stepwise methods invariably select a first question with minimal expected path length. The proof is in Appendix A.*

Theorems 5 and 6 hold for relatively simple Person Game environments, specifically those with fewer than eight items or fewer than three constraint questions. For more complex scenarios, we used computer simulations to estimate how frequently environments occur in which the stepwise methods fail to identify the best first question. Because adding a single additional item roughly doubles the computation time for determining the objectively best first question through dynamic programming (Appendix B), we simulated environments of up to 18 items.

#### 6.4.2 Computer simulations

It is implausible, to say the least, that people would explicitly use dynamic programming to identify the very best question tree in medium or large environments. However, what if perceptible environmental characteristics could give a clue to the tendency of the stepwise methods to be suboptimal? To explore this, we tested about 1 million randomly generated Person Game environments, manipulating several environmental characteristics (see Appendix C). The environments that we generated went from the simplest and least plausible (e.g., independent and identically distributed features) to environments

with more complexity, such as inter-feature dependencies that might be found in natural environments. The specific environment characteristics manipulated included the number of items, the number of constraint features, and the average density for constraint features (i.e., the expected number of items possessing a particular feature). In addition, features were generated in four different ways that crossed two binary factors. First, features either were generated independently or tended to be correlated with each other. Second, the generating density was either the same for all features or variable. In each randomly generated environment, the simulation checked whether the stepwise methods would select the globally optimal (i.e., most efficient) first question.

### 6.4.3 Simulation results

We discuss highlights here; see Appendix C and Figures C1-C4 for full methods and results. Figure 5 shows the proportion of environments in which the stepwise methods failed to identify the best first question, under each of the four methods for generating features. Figure 5a plots this proportion as a function of the number of items (varied from 6 to 18, in steps of 2), and Figure 5b plots it as a function of the number of constraint questions in the game (all values from 2 to 10, and 20). The results provide several insights. First, the simulations illustrate Theorems 5 and 6: if there were fewer than three constraint questions, or fewer than eight items, the stepwise methods invariably identified the best first question. Moreover, the tendency of the stepwise methods to be suboptimal decreased again when there were a large number of constraint questions (Figure 5b). Second, when considering somewhat more complex environments, the suboptimality of the stepwise strategies is not limited to the specific example of Table 1: across all conditions with 8 to 18 items (persons), the stepwise methods failed to identify the best first question in multiple environments. Third, the simulations provide clues about which environmental characteristics promote or impede the success of the stepwise methods. Suboptimality of the stepwise methods was the most frequent when features tended to be correlated with each other, and when they were diverse in

their densities. In the most difficult condition for the stepwise methods, viz. with 18 items and features that are correlated and diverse in their density, these methods led to selecting a suboptimal first question in 10% of random Person Game environments generated. Fourth, although the stepwise methods might be deemed to achieve “reasonable” performance, their tendency to be suboptimal increased markedly with the number of items. Thus, it seems likely that in task environments with larger numbers of items (e.g.,  $n = 100$ ), the stepwise methods would be suboptimal much more frequently than in the present simulations.

#### **6.4.4 Pathways for empirical research**

Our simulations demonstrate that the tendency of the stepwise methods to be suboptimal varies systematically with specific environmental characteristics. The stepwise methods did better if features were generated independently of each other, if features had about the same density, and if there were a large number of constraint questions. These observations lead to several testable hypotheses regarding human search behavior. First, there is the question of whether people have second-order knowledge of when the stepwise methods are likely to be effective. If so, then the environmental characteristics identified in the simulations here should be predictive of people’s tendencies to follow stepwise methods. For example, one could create two environments in which stepwise methods identify a suboptimal first question, one with a small number of items and a large number of independent features of homogenous density, and the other with a large number of items and a small number of correlated features with diverse densities. To the extent that people utilize these environmental meta-features to guide strategy selection, they should be more likely to follow stepwise methods, and hence less likely to select the optimal first question, in the former environment than the latter. Second, there is the related question of whether people have explicit (verbal) knowledge of when stepwise methods are likely to be adequate. This question could be addressed by asking people to report their beliefs about which kinds of environments (or environmental characteristics) predict good performance of the split-half

strategy. Third, there is the question of whether this sort of second-order knowledge is learnable. That is, can people learn through repeated experience when they can rely on stepwise methods and when they should plan ahead? This question could be addressed in longer-term learning studies, wherein subjects are exposed to many environments varying in their characteristics and given feedback regarding the outcomes of their choices.

### 7. Planning ahead in the Person Game: Multi-step OED models

Although OED models are traditionally defined for single queries (i.e. implemented in a stepwise fashion), they can be adapted for planning  $k$  steps ahead (where  $k > 1$ ). In the case of planning two steps ahead, a searcher may consider simultaneously which feature to query first, and which feature will be queried next, contingent on the outcome of the first query. The resulting 2-step tree could be evaluated using the same OED procedures that are applied to single queries, for example by calculating the expected information gain that would result once one reaches the end of the tree. Put differently, any multistep decision tree is formally equivalent to a single many-valued query, with possible outcomes corresponding to the leaf nodes of the tree. The informational utility of any tree (e.g., its expected information gain or probability gain) is equal to that of the corresponding many-valued query. We refer to such a tree evaluation procedure as a *multi-step OED* model. Such an approach is less myopic than one that evaluates only the immediate next query, but the planning horizon for 2-step OED is still limited because only the next 2 queries are considered. In general,  $k$ -step strategies can be defined as strategies that evaluate trees of depth  $k$  with respect to some OED criterion, with the criterion being calculated for the tree as a whole rather than for each question individually.

For convenience, we also refer to  $k$ -step OED strategies applied to an arbitrary question tree (of depth greater than  $k$ ), by evaluating the first  $k$  steps of the tree. In Figure 3, the top row shows several complete trees for the example Person Game environment of Table 1 (including the stepwise-optimal and globally optimal trees discussed in the previous section).

The middle row shows the corresponding subtrees for depth  $k = 2$ , and the bottom row shows the subtrees for depth  $k = 3$ . For example, 3-step probability gain applied to the stepwise optimal tree (Figure 3a, top) is defined as the probability gain of the subtree obtained by removing all queries beyond step 3 (i.e., of the tree in Figure 3a, bottom).

Multistep OED strategies can lead to superior performance as compared to the corresponding stepwise OED strategies. Table 2 compares  $k$ -step impact, information gain, and probability gain across the four main trees in Figure 3 for different values of  $k$ . Stepwise information gain and stepwise impact ( $k = 1$ ) both prefer query  $Q_1$ , which forms the root of the stepwise-optimal tree (Figure 3a, top row) over  $Q_2$ , which forms the root of the globally optimal tree (Figure 3b, top row). However, when planning two ( $k = 2$ ) or three ( $k = 3$ ) steps ahead, information gain and impact both follow the globally optimal tree; that is, considering two or more steps at once allows these strategies to identify the optimal first query. Planning  $k = 3$  steps ahead is required, in this example, for probability gain to uniquely identify the globally optimal first query.

### 7.1 Multistep probability gain

The next three subsections focus on multistep probability gain. Multistep probability gain is interesting because (i) stepwise probability gain is indifferent among informative questions; but (ii) two-step ( $k = 2$ ) probability gain differentiates among informative questions (Table 2) in a way that corresponds to intuitive distinctions in early literature on the Person Game (explained below); and (iii) as we show here, there are simple strategies for evaluating trees'  $k$ -step (multistep) probability gain, which may provide the basis for developing candidate psychological models.

In any Person Game environment, there always exist *exhaustive question trees*, meaning trees that will identify the true hypothesis (item or person) with certainty, if there is no limit on the number of questions that can be asked. (Recall that in the Person Game, every item has a unique identifying name feature.) Indeed, the stated goal of the game is to continue until the

target is known. All exhaustive trees have the same probability gain,  $1-1/n$ , which equals .875 for our example environment (see Table 2 and Figure 3).

From a psychological standpoint, cognitive (for example, working memory) constraints may sharply limit the planning time horizon. What probability gain can be achieved on the Person Game if only a limited number of questions can be asked, or if a searcher plans only a small number of questions in advance, such that the decision trees under consideration are non-exhaustive (i.e., have leaf nodes containing more than one hypothesis)? To illustrate, the 2-step tree obtained from the globally optimal tree (Figure 3b, middle row) has four leaf nodes respectively containing 1, 2, 2, and 3 hypotheses. One-step probability gain, which considers only the first question in each tree, does not distinguish among the trees in Figure 3; however, 2- and 3-step probability gain do make distinctions among these trees (Table 2). Specifically, the trees that begin with questions that leave multiple hypotheses consistent with each answer yield greater 2-step probability gain than do the trees that start with a hypothesis-scanning question (i.e., that isolate a single hypothesis from the rest). Thus at depth  $k=2$  (middle row of Figure 3), probability gain prefers a tree that does not start with a name (hypothesis-scanning) question, although it does not differentiate between the stepwise- and globally optimal trees. However, when the trees are evaluated to a depth of  $k=3$  (bottom row of Figure 3), the globally optimal tree has higher probability gain than the stepwise-optimal tree. Thus, when planning three queries ahead, probability gain correctly recognizes the most efficient first query.

Why do the trees differ in their probability gains when planning multiple steps ahead? It turns out that a tree's probability gain is a function of the number of its (nonempty) leaf nodes:

**Theorem 7:** *In the Person Game, the expected probability gain of any question strategy (tree  $T$ ) is given by  $eu_{PG}(T) = (r-1)/n$ , where  $r$  is the number of (nonempty) leaf nodes of the tree, and there are  $n$  persons (items). The proof is given in Appendix A.*



When evaluating the  $k$ -step probability gain for a tree of depth greater than  $k$ ,  $r$  is taken to be the number of leaf nodes of the subtree defined by the tree's first  $k$  steps. For example, 3-step probability gain for the stepwise-optimal tree (Figure 3a, top) equals  $(6-1)/8 = .625$ , because the tree's 3-step subtree (Figure 3a, bottom) has 6 leaf nodes.

The intuitive explanation for Theorem 7 is that, regardless of which leaf node is reached, the player can select one target as his or her guess, from among the targets at that leaf node. Imagine these  $r$  candidate guesses are designated in advance (the timing has no impact on expected performance). Then a priori, the player's guess will be correct if and only if the selected target is among this set, meaning the guess will end up being correct with probability  $r/n$ .

Therefore, multistep probability gain embodies the goal of avoiding "dead ends" on a question tree. Consider the two-step question tree beginning with a name question in Figure 3c, middle row. The possibility of uniquely identifying the target after only one question (if the target is  $h_1$ ) is actually a disadvantage of this tree, because it squanders the opportunity to ask another question at Step 2. That is, this path is a premature dead end, and it leads to the tree having only three leaf nodes at depth 2, rather than four as the trees in Figures 3a and 3b do.

In general, the maximum number of leaves for a tree of depth  $k$  is  $2^k$ . If this maximum is achieved, we refer to the tree as *nondegenerate* or *complete*. Likewise, we refer to an arbitrary tree as *complete at depth  $k$*  if it has  $2^k$  nodes at depth  $k$ , i.e., if the subtree defined by its first  $k$  steps is nondegenerate. If a tree is complete at depth  $k$ , then it necessarily has maximal  $k$ -step probability gain. A tree can be complete at depth  $k$  only if  $2^k \leq n$ , and complete exhaustive trees can exist only if the number of persons  $n$  is a power of 2. However, even at shallower depths  $k$  where  $2^k < n$ , complete trees might not exist in any particular Person Game. For example, it is possible when constructing a tree to reach a node at which all informative questions differentiate only a single member from the remaining hypothesis set. The child

node containing that single hypothesis will necessarily be a leaf node that cannot be further split. The goal of maximizing  $k$ -step probability gain is thus equivalent to that of minimizing the number of degeneracies. In the Table 1 environment, trees starting with name questions (Figures 3c & 3d) are degenerate at step 2. The stepwise-optimal tree (Figure 3a), as well as any other tree starting with query  $Q_1$ , is doubly degenerate at step 3, because only name questions remain after  $Q_1$  is answered. The globally optimal tree (Figure 3b) has only a single degeneracy at step 3 making it optimal for 3-step probability gain.

## 7.2 Pathways for empirical research

The extensions derived here raise a number of new empirical issues. The first issue is which, if any, of the multistep OED models describe people's preferences in such a task. Multistep probability gain, taken as a predictive model for human behavior, makes the strong prediction that a question tree's utility is determined fully by its number of nonempty leaf nodes. Multistep information gain and impact, however, predict that the distribution of hypotheses among these nodes also matters. For example, 2-step information gain prefers the tree in Figure 3b to that in Figure 3a, because the former has a more even distribution of hypotheses among the nodes at Depth 2 (1,2,2,3 vs. 1,1,3,3), whereas 2-step probability is indifferent because both trees have four nodes at that depth.

Second, do people spontaneously notice, or can people learn through experience, that on Person Game tasks, nondegenerate trees (e.g., at depth  $k=2$  or  $k=3$ ) have higher probability gain than degenerate trees? That is, do people recognize that dead ends on a tree waste the opportunity for further questions? Would explicit visualization of the tree structure help people notice this property of multistep question strategies? What kind of visualizations would people spontaneously employ, and could people be trained to create helpful tree-like visualizations on these tasks?

Third, do preferences differ depending on how the task is carried out? As noted above, any multistep decision tree is formally equivalent to a single many-valued query. However, it

is not known whether equivalent empirical results would be obtained if a task were framed in terms of selecting a decision tree, as opposed to the traditional sequential question asking framing.

### **7.3 Mathematical implications of the constraint vs. name question distinction**

The developmental psychological literature on the 20-Questions game (e.g., Mosher & Hornsby, 1966; Denney & Denney, 1973) deems constraint questions to be more useful than hypothesis-scanning (name) questions. Theorem 7 shows how the constraint-versus-name question distinction has precise mathematical implications, in the context of two-step probability gain. Consider Person Games with four or more items, which is necessary for constraint questions to exist, and suppose that the goal is to achieve the highest classification accuracy possible, given that up to two questions can be asked. Thus, 2-step probability gain exactly captures the player's objective. If a tree starts with a name question (e.g., Figure 3c, d), then it cannot include a follow-up question on one of its branches in the second step (because in the case of an affirmative answer only the target item remains). Therefore, regardless of the question chosen for Step 2 on the other branch, the tree will have exactly 3 leaf nodes and hence yield accuracy  $3/n$  and probability gain  $2/n$ . In contrast, if a tree starts with a constraint question, then follow-up questions are possible on both branches, and thus the tree can have 4 leaf nodes and hence yield accuracy  $4/n$  and probability gain  $3/n$ . Thus, the intuitive distinction between constraint questions and name questions has a precise mathematical implication, in the context of the goal of classification accuracy in the two-step Person Game. If you start with a constraint question, you are guaranteed to have at least one further informative question to ask, no matter the answer to your first question, and therefore your expected accuracy after two steps is greater than if you start with a name question.

### **7.4 Constraint questions can be inferior to name questions**

Although the results of the previous subsection provide one formal understanding of the value of constraint questions, this result is particular to the goal of expected accuracy after

two queries. Under the goal of uniquely identifying the target with the minimal number of questions, constraint questions are not in general better than name questions on the Person Game. Consider the Person Game environment of Table 3 and Figure 6, with 10 items, 3 constraint features, and the 10 individuating ("name") features. If you start with constraint question  $Q_3$ , and choose optimally thereafter, a total of 4.7 questions will be required on average. Most of the name questions ( $Q_4$  through  $Q_9$ ), however, will require only 4.4 questions on average. As before, the difference arises from the kinds of follow-up questions available: After asking  $Q_3$ , all remaining questions are (effectively) name questions, but after asking any of  $Q_4$ – $Q_9$  there are two more constraint questions that can be asked. Whether people would be sensitive to these distinctions is not known, and would be another way to address whether people focus on stepwise properties of queries (here, the name vs. constraint distinction) or are sensitive to multistep considerations.

### 7.5 Summary and implications

Suppose the goal is to maximize accuracy after many questions have been asked, but at each point in the game, it is possible to plan only a small number of steps in advance. That is, the objective is  $m$ -step probability gain, but the agent has access only to  $k$ -step methods ( $k < m$ ). Our results show, somewhat paradoxically, that probability gain is not the most effective method in this case. If one's *goal* is to maximize long-run accuracy, it does not follow that stepwise probability gain is necessarily the best stepwise *strategy* ( $k = 1$ ). Unlike stepwise impact and information gain, stepwise probability gain does not distinguish among informative questions (Theorem 4). Two-step probability gain ( $k = 2$ ) is a more reasonable strategy, corresponding to the constraint- vs. name- question distinction in the developmental literature, but it still is less sensitive than impact and information gain, which differentiate among constraint questions according to their splithaleness. Moreover, it is also possible for a name question to be better than a constraint question: in the example of Table 3, starting with any of  $Q_4$ – $Q_9$  yields greater  $m$ -step probability gain than starting with  $Q_3$ , for  $5 \leq m \leq 8$ . On

the other hand, as the depth of lookahead ( $k$ ) increases, probability gain becomes more fine-grained in its distinctions among available queries' values. Furthermore, a tree's  $k$ -step probability gain can be calculated in a straightforward way, simply by counting its leaf nodes (Theorem 7). This operation is potentially much simpler, psychologically, than calculating other multistep OED metrics, which depend on quantities like expected posterior entropy. As discussed in Section 7.2, these opposing considerations suggest paths for empirical research.

## 8. The Generalized Person Game: Unequal priors

In this section we discuss the implications of unequal priors in what we call the *Generalized Person Game*. This game is identical to the Person Game, except that the prior over the hypotheses (items or people) is not necessarily uniform (i.e. the prior probability of each hypothesis  $P(h_j)$  is not necessarily  $1/n$ ). The other characteristics of the Person Game, for instance that questions are binary and likelihoods are deterministic, still hold. We first give a theoretical result on how to calculate the classification accuracy (and probability gain) of a  $k$ -step question tree in the Generalized Person Game. We then use a simple example to illustrate two points: (1) the split-half heuristic does not in general correspond to maximizing stepwise information gain, and (2) the split-half heuristic can lead to better performance than stepwise information gain in the Generalized Person Game.

### 8.1 Maximizing accuracy in a $k$ -step Generalized Person Game

What if the goal is to maximize accuracy, given that up to  $k$  questions can be asked, in the Generalized Person Game? Here we generalize the result from Theorem 7 on the probability gain of a regular Person Game tree at depth  $k$ , to the Generalized Person Game, with arbitrary priors.

**Theorem 8:** *in the Generalized Person Game, the probability gain of any question strategy (tree  $T$ ) is given by  $eu_{PG}(T) = \sum_{j=1}^r P(h_j^*) - \max_{1 \leq j \leq r} P(h_j^*)$ , where the terminal*

nodes of the tree are indexed by  $j=1, 2, \dots, r$ , and  $h_j^*$  is a hypothesis in node  $j$  that has highest prior probability. The proof is given in Appendix A.

Extending the reasoning given after Theorem 7, for each leaf node ( $j$ ) the searcher can choose one hypothesis ( $h_j^*$ ) as his or her guess if the query process terminates at that node. When the prior is heterogeneous, the chosen hypothesis will (optimally) have the greatest prior probability among the hypotheses at that node. A priori, if the sampled target is any of the  $h_j^*$ , then the guess will end up being correct. Therefore, to calculate the classification accuracy of a tree, one needs only to sum the prior probabilities of all the  $h_j^*$ . Furthermore, the classification accuracy prior to asking any questions is determined by the hypothesis with the greatest prior probability, which is necessarily one of the  $h_j^*$ . Therefore the probability gain of the tree is  $eu_{PG}(T) = \sum_{j=1}^r P(h_j^*) - \max_{j=1:r} P(h_j^*)$ . In the special case of uniform priors,  $h_j^*$  is arbitrary at each leaf node and  $P(h_j^*) = 1/n$ . Thus expected accuracy is  $\sum_{j=1}^r \frac{1}{n} = \frac{r}{n}$ , and the expected probability gain is  $r/n - 1/n = (r-1)/n$ . Thus, Theorem 7 can easily be seen to be a special case of Theorem 8.

## **8.2 In the Generalized Person Game, the split-half heuristic corresponds neither to maximum question entropy, nor to information gain, nor to impact**

Consider a Generalized Person Game in which Mickey Mouse is one among a total of eight items, Mickey is chosen with probability  $\frac{1}{2}$ , and the other seven items each with probability  $1/14$ . Suppose that the only constraint question is gender (male or female), and that there are three other male faces and four female faces. The probability of a positive answer to the “Is it Mickey?” question is  $\frac{1}{2}$ . That question is the maximum entropy question, and therefore also the maximum information gain question, because likelihoods are deterministic in the Generalized Person Game (i.e., Theorem 1 applies). The split-half heuristic asks about gender, which has lower information gain. Thus, the correspondence of

the split-half heuristic to maximum stepwise information gain (Theorem 3) does not hold in the Generalized Person Game.

The fact that the split-half heuristic begins with a question that has less than maximal information gain does not, however, entail that the split-half heuristic is inferior to information gain. Consider the depth  $k=2$  trees, in the Mickey Mouse example, that would be selected by stepwise information gain and by the split-half heuristic (see Figure 7). The stepwise information gain tree (Figure 7a) starts with the maximum information gain question (“Is it Mickey?”). If yes, the true face is known to be Mickey, so no further questions are asked. If the answer is no, the next question is gender. The split-half heuristic starts by querying gender. Regardless of the answer, the next question is a name question chosen at random among the four remaining items. Note that if the gender is male, the split-half heuristic might query Mickey on Step 2 (Figure 7b), or it might query a different male face (Figure 7c), because the heuristic is insensitive to base rates. Either way, the two-step probability gain of the resulting split-half tree is  $3/14$ , as compared to  $2/14$  for the stepwise information gain tree. This result follows immediately from Theorem 8; intuitively it holds because only the split-half tree guarantees the possibility of an informative second question.

### 8.3 Empirical tests in the Generalized Person Game

The results of the previous two subsections suggest a number of possible experimental manipulations. Theorem 8 entails that 1-step probability gain is maximized by any question that splits (i.e., yields opposing answers for) the top two hypotheses. Imagine a Generalized Person Game with a single most-probable hypothesis (Mickey, with probability 0.3), a second-most-probable hypothesis (Minnie, with probability 0.2), and several additional hypotheses with equal probability (the seven dwarves, each with probability  $1/14$ ). People could be incentivized to maximize one-step probability gain by being allowed to ask a single question for free, and then given a prize if they correctly guess the true item. Under these conditions, “Is it Mickey?” and “Is it Minnie?” both yield maximal expected accuracy of .5

(in either case,  $\{h_j^*\} = \{\text{Mickey, Minnie}\}$ ). In contrast, "Is it a Mouse?" yields maximum stepwise information gain (1 bit), and a question that divides four of the dwarves from the other items would be maximally splithalfy, but both of these are suboptimal according to stepwise probability gain (expected accuracy .37). Therefore it would be of interest to see whether people can adapt to the demands of the task and select a question maximizing one-step probability gain over those maximizing information gain or splithalfiness, and further whether people would recognize that "Is it Minnie?" is as effective a query as "Is it Mickey?", even though Minnie has lower prior probability than Mickey. Further manipulations would be to use payoff structures that explicitly relate to multistep probability gain, for instance by allowing up to two questions to be asked for free, and then offering a prize just in case the correct item is guessed after the second question. In this case, an optimal two-step tree is any that splits the top four hypotheses, and it would be of interest to test whether people would choose accordingly.

Ruggeri, Sim and Xu (2017) found evidence that children ages 3-5 years were sensitive to the prior probabilities of hypotheses, asking questions to maximize information gain in instances in which it contradicted the split-half heuristic. Although Ruggeri et al. did not use a multistep procedure and did not attempt to differentiate stepwise versus global optimality, their results provide a tantalizing foundation for future empirical research.

### **9. Stepwise procedures can be suboptimal given only two hypotheses**

How general is the finding that what is optimal for one-shot tasks may be suboptimal in the long run? How simple can an environment be and still exhibit this conflict? Theorem 6 shows that, if there are only two constraint questions in a Person Game, stepwise and global optimality necessarily agree. However, we show here that this restriction does not hold in every task. In particular, if likelihoods are allowed to be nondeterministic, then stepwise and global optimality can diverge even in scenarios with just two features and two hypotheses.



In this section we consider a two-category, binary-feature scenario (“Planet Vuma”) that is one of the most frequently used tasks in psychological research on human information selection (Meder & Nelson, 2012; Nelson, 2005a; Nelson et al., 2010; Rusconi et al., 2014; Skov & Sherman, 1986; Slowiaczek et al., 1992; Wu et al., 2017). In the Planet Vuma scenario, the task is to categorize a fictitious alien into one of two species (hypotheses  $h_1$  and  $h_2$ ) by querying their (binary) features. In versions of the task used to date, a single question can be asked before making a classification decision (e.g., “Does the alien wear a hula hoop?”), although here we extend the task to cases where multiple questions can be asked. In addition, this scenario differs from the Person Game by having arbitrary (nondeterministic) likelihoods and arbitrary priors. In other words, knowledge of the species of creature would not in general make it possible to know whether or not the creature has a particular feature, and hence  $\text{ent}(Q|H) > 0$ . One consequence is that on Planet Vuma, the maximum entropy question strategy does not necessarily correspond to maximum information gain, and Theorem 1 does not apply. Furthermore, it is not in general possible (no matter how many questions are asked) to achieve certainty about the true category. Therefore, we focus on the goal of achieving maximum possible classification accuracy, given a fixed number of queries (i.e.,  $k$ -step probability gain).

The research question in these papers is which properties of a query cause people to consider it informative for discriminating among the hypotheses (categories). In relatively simple scenarios, people typically have good intuitions about what makes features useful (Nelson, 2005a; Skov & Sherman, 1986; Slowiaczek et al., 1992). However, the way in which information about the statistical structure of the environment is conveyed strongly matters, with human search behavior systematically varying as a function of presentation format. Most experiments using the Planet Vuma scenario have presented prior probabilities and feature likelihoods with words and numbers; under these conditions behavior is often indistinguishable from chance. In contrast, Nelson et al. (2010) found that if people learn

environmental probabilities through many trials of experience, with feedback, they are very sensitive to the probability gain of each question and preferentially select questions with high expected probability gain (see Crupi et al., 2018, for a discussion of alternate models).

In parallel to our discussion of the split-half heuristic on the Person Game, it is interesting that a heuristic strategy has also been reported on the Planet Vuma task, when probability information is presented with numeric prior probabilities and feature likelihoods (Nelson, 2005a; Slowiaczek et al., 1992). The *likelihood difference heuristic* selects queries by evaluating the absolute difference in feature likelihoods between the hypotheses. Consider a binary feature  $A = \{a_1, a_2\}$ . Its likelihood difference,  $\text{LikDiff}(A)$ , is equal to  $P(a_1 | h_1) - P(a_1 | h_2)$ . (Note that  $P(a_1 | h_1) - P(a_1 | h_2) = P(a_2 | h_1) - P(a_2 | h_2)$ , so it does not matter whether the likelihood difference is defined in terms of feature value  $a_1$  or  $a_2$ .) If there are multiple features, the heuristic queries the feature with highest likelihood difference. It turns out that the question (feature) with highest likelihood difference is also the question with highest expected impact (proof in Nelson, 2005a, Footnote 2). This example provides another case in which a simple heuristic implements an OED model. In the special case where  $P(h_1) = P(h_2) = .5$ , impact and probability gain are also identical, and the likelihood difference heuristic also selects the question with highest probability gain.

### 9.1 Stepwise vs. multistep strategies given binary hypotheses

Here we consider an extended version of the Planet Vuma task that allows for multiple queries on each trial. In this extended variant of the task, a sample with one or more biological specimens—all from the same, but unknown, location—has been obtained, and the task is to infer whether the sample is from location  $h_1$  or location  $h_2$ . There are two genes, each of which can take one of two forms: gene  $A$  takes value  $a_1$  or  $a_2$ , and gene  $E$  takes value  $e_1$  or  $e_2$  (the organisms are haploid). The distribution of the two forms of each gene at each location is known. Neither gene perfectly predicts the location of origin of the specimens. From a mathematical standpoint, the genes are class-conditionally-independent (Domingos &

Pazzani, 1997; Jarecki, Meder & Nelson, 2018) and identically distributed across specimens. Each specimen can be given only a single test, due to their small size and limited amount of genetic material. Thus each query is a measurement of either gene A or gene E, and the total number of possible queries equals the number of specimens in the sample.

Suppose it is known that for any particular specimen,  $P(h_1) = .70$ ,  $P(h_2) = .30$ ,  $P(a_1 | h_1) = .04$ ,  $P(a_1 | h_2) = .38$ ,  $P(e_1 | h_1) = .57$ , and  $P(e_1 | h_2) = 0$  (Figure 8a). These environmental probabilities are from Nelson et al. (2010, Supplemental Material) and were designed to maximally differentiate the predictions of stepwise information gain and stepwise probability gain. Suppose that there is just one specimen in the sample, so only a single test can be conducted. Which test is more useful if the goal is to classify the specimen's location of origin? Figure 8b illustrates the search problem. If only a single test can be conducted, Test *A* has probability gain of .086, information gain of .132, and impact of .286. Test *E* has zero probability gain, information gain .280, and impact .479. Thus, if the goal is to maximize classification accuracy, Test *A* is better. If the goal is to maximize information gain or impact, Test *E* is better.  $\text{LikDiff}(A) = .34$ , whereas  $\text{LikDiff}(E) = .57$ , so the likelihood difference heuristic would select Test *E*.

What if the sample contains not one but two specimens, from the same location, and it is possible to conduct one test on the first specimen, and contingent on the result of that test, to decide which test to conduct on the second specimen? Should Test *A* be conducted on both specimens? Should Test *E* be conducted under some circumstances, even if the goal is to maximize accuracy? There are  $2^3 = 8$  possible binary decision trees (question strategies) that can be formed, according to the choice of first test and the choice of the second test conditional on the two possible outcomes of the first test. However, if Test *E* is conducted on the first specimen, and the result is  $e_1$ , the true hypothesis is known to be  $h_1$  (Figure 8b). Thus we can prune the trees where the first result is  $e_1$ , leaving six trees (Table 4).

What happens if tests are chosen by using probability gain in a stepwise manner? In this case, Test  $A$  is used for the first specimen. If the result is  $a_1$ , Test  $E$  is used on the second specimen; if the result is  $a_2$ , the  $A$  test is repeated on the second specimen. The stepwise probability gain tree (Figure 8c) uses exactly two questions; it has two-step probability gain of .146, two-step information gain of .274, and two-step impact of .470.

What happens if information gain is used in a stepwise manner? In this case, Test  $E$  is used for the first specimen. If the result is  $e_1$ , the true hypothesis is known to be  $h_1$ , and no further tests are conducted; if the result is  $e_2$ , another  $E$  test is conducted. This tree is depicted in Figure 8d; the same tree results if impact is used in a stepwise manner. It may not be surprising that the stepwise information gain tree has higher two-step information gain (.502) than the stepwise probability gain tree (.274). Interestingly, the stepwise information gain tree also has higher two-step probability gain than the stepwise probability gain tree (.171 vs. .146)!

Consideration of all six possible trees (Table 4) shows that the stepwise information gain tree is in fact globally optimal with respect to the goals of maximizing classification accuracy, minimizing posterior entropy, and maximizing impact. This tree also ties for requiring the fewest expected number of tests. The likelihood difference heuristic, which always chooses the test with highest stepwise impact, also selects this globally optimal tree.

In summary, in a paradigmatic case from the psychological literature, if you plan just two steps ahead (rather than one), the relative usefulness of the two questions as a *first* query can actually reverse. Furthermore, for this example scenario, stepwise strategies of maximizing impact or information gain identified the tree that maximizes multistep probability gain, whereas stepwise probability gain did not. Are people sensitive to this discrepancy between stepwise and multistep considerations in this context? Is the better multistep performance of information gain responsible for the "information bias" observed in

a related task (Baron, Beattie & Hershey, 1988)? These are important issues for future empirical research.

## **9.2 Simulation of performance of stepwise OED strategies in two-step, two-category, binary-feature tasks**

Is the above example unusual? To address this, we tested the tendency of stepwise probability gain, impact, and information gain to choose as a first question the feature that the highest 2-step probability gain tree begins with. The simulation looped through prior probability  $P(h_1)$  values of [.01, .05, .10, ..., .95, .99] and likelihood  $P(a_1|h_1)$  values of [0, .05, ... 1.0]. For each of the combinations of  $P(h_1)$  and  $P(a_1|h_1)$  values, 10000 random scenarios were generated, with the other likelihoods, namely  $P(a_1|h_2)$ ,  $P(e_1|h_1)$ , and  $P(e_1|h_2)$  chosen independently from a uniform(0,1) distribution. The globally optimal tree for each scenario was defined as the tree with highest expected two-step probability gain. The simulation checked, for each scenario, whether stepwise probability gain, impact, and information gain identified the globally optimal tree (i.e., the 2-step tree with maximal 2-step probability gain). Across the approximately 4.4 million random scenarios that were simulated, stepwise information gain failed to identify a unique best query in 3.1% of the scenarios, stepwise impact in 5.1% of the scenarios, and stepwise probability gain in 7.0% of the scenarios. The fact that stepwise probability gain is not the best stepwise strategy with respect to longer-run probability gain is interesting, and parallels the results from the Person Game simulations.

## **9.3 Experimental tests**

In future work, it will be important to establish the circumstances under which people can identify the best first question in sequential versions of the Planet Vuma task. The example discussed in Section 9.1 (Figure 8 and Table 4) is one specific environment in which stepwise probability gain conflicts with stepwise information gain and impact as well as global optimality. Environments that arose in the simulations in Section 9.2 also exhibit other patterns of agreement and conflict among these four measures. Systematic investigation of

human behavior in these environments should be able to determine what factors drive information selection in multistep tasks using nondeterministic queries. Moderation by the format of presentation could also be investigated: In addition to the standard probability format, methods for presenting the probabilistic information could involve experience-based learning (e.g., Nelson et al., 2010), or helpful numeric or graphical formats (Wu, Meder, Filimon & Nelson, 2017).

## 10. General Discussion

Much work on active information gathering is based on the idea that stepwise OED models should be used to determine which test to conduct. Our results challenge the preeminence of stepwise OED models in two complementary ways.

First, simple heuristics can exactly implement stepwise OED models, without directly computing the informational utility functions on which the OED models are based, in many situations. Thus, rather than there being a fine line between "optimal" and "heuristic" models, there may in some cases be no dividing line at all. In the Person Game, the split-half heuristic identifies the question with maximum stepwise impact (Theorem 2) and information gain (Theorem 3). In the two-category, binary-feature ("Planet Vuma"; Skov & Sherman, 1986) scenarios, the likelihood difference heuristic corresponds to the OED model of impact maximization (Nelson, 2005a). These results provide pathways for developing psychological theory that integrates normative and descriptive considerations.

Second, much work using OED models is based on the incorrect idea that stepwise strategies—in particular, stepwise information gain—are maximally efficient for sequential search. Stepwise OED models, which are often considered rational benchmarks for information search in one-shot tasks, are not necessarily optimal for planning a sequence of queries. We have shown via a simple example (Table 1 and Figure 1) and extensive simulations that maximizing stepwise information gain does not in general maximize long-run

efficiency in the Person Game. Surprisingly, although the goal in this task is to achieve perfect accuracy in the most efficient way, stepwise probability gain (which maximizes stepwise increase in accuracy) is indifferent among non-redundant questions in the Person Game (Theorem 4). We showed via examples and simulations that stepwise probability gain does not maximize two-step probability gain in the popular two-category, binary-feature (Planet Vuma) information-acquisition scenario either. (Further simulations show that stepwise information gain does not always maximize two-step information gain, either.) Remarkably, which initial question is the best from the perspective of probability gain can actually reverse if two questions, rather than a single question, can be asked.

The Person Game and Planet Vuma are key tasks in psychological information acquisition literature. Moreover, they are simple in complementary ways: the Person Game has many hypotheses but binary, deterministic features, whereas Planet Vuma has arbitrary likelihoods (i.e., nondeterministic features) but only two hypotheses and, in the scenarios considered here, only two features. Stepwise OED methods can fail in both of these tasks. Moreover, Person Game simulations show that the stepwise methods fail more frequently as the toy games get larger (with 18 items, rather than 8 items), or the feature construction methods get more naturalistically plausible (e.g., with some simple inter-feature dependencies and heterogeneity in feature densities). It is therefore plausible that stepwise methods are suboptimal in many or even most realistic scenarios.

Our final contribution is the analysis of multistep models for question selection. We showed that probability gain, which is indifferent among informative questions when applied in a stepwise way, is not indifferent when planning more than one step ahead in the Person Game (Theorems 7 and 8). Two-step probability gain gives a mathematical characterization of how constraint and name questions differ in their information value: Trees that start with a name question have two-step probability gain  $2/n$ , whereas trees that start with a constraint question have two-step probability gain  $3/n$ , in any Person Game with  $n$  items. Surprisingly,

however, constraint questions are not always better from the perspective of long-run efficiency (Table 3). For any fixed number of persons (hypotheses), a Person Game tree's probability gain is a linear function of that tree's number of leaf nodes (Theorem 7). Theorem 8 shows that even in the Generalized Person Game, with arbitrary prior probabilities, a fairly straightforward strategy can be used to calculate a tree's probability gain. In fact, Theorem 8 also applies to other 20 questions tasks (Figure 4), including the Person Game, the Number Game, and the Generalized Number Game.

### 10.1 Stepwise strategies to see the future

Given that maximizing expected accuracy seems to approximate people's goals on classification tasks in which just one query can be conducted before a decision can be made (Nelson et al., 2010), it is especially interesting that stepwise probability gain is not even the best *stepwise* strategy to efficiently achieve high accuracy over multiple steps, in the Person Game and Planet Vuma. Similar theoretical results have been found in at least two other tasks that are highly relevant to cognitive science, as we discuss next.

Planning active interventions in causal learning is one domain in which the combinatorics of planning in advance strongly suggest looking for reasonable myopic strategies. Bramley et al. (2015) explored the use of stepwise OED strategies in a causal learning task with dozens of possible hypotheses, possible interventions in the causal system, and possible outcomes of each intervention. The learner's task was to figure out which of 25 acyclic graph structures connected three nodes. For instance, the nodes could be connected by a chain (e.g.,  $A \rightarrow B \rightarrow C$ ), or could have a common-cause structure (e.g.,  $B \leftarrow A \rightarrow C$ ), or could have one or more nodes unconnected, and so on. Interventions included all possible combinations of clamping individual nodes on or off, and observing the activations or lack thereof of the other nodes (Bramley et al., 2015, Fig. 4). Human subjects' queries had high information content. Of particular theoretical interest, however, is the relationship among the different stepwise methods, independent of the human data, in a simulation study (Bramley et



al., 2015, Figure 9). Stepwise probability gain, stepwise information gain, and stepwise utility gain were tested. Utility gain was based on the actual payoff structure of the task, namely earning one point per correctly identified link among the three nodes in each problem. Assuming that a stepwise strategy must be used, is it better to select queries according to stepwise improvement in utility gain, probability gain, or information gain? Remarkably, after approximately six queries had been selected with each stepwise method, stepwise information gain obtained the highest utility score on the task, even better than stepwise expected utility gain.

Najemnik and Geisler (2005) obtained similar findings, in a very different domain involving eye movements for perceptual visual object detection. In Najemnik and Geisler's task, a visual object must be located in a field of noise and many distractor objects. The object being searched for is placed at one of several dozen different possible locations. Thus, the task is formally similar to a probabilistic classification task (the target is at exactly one location, just like an object is exactly one category), where eye movements are analogous to explicit questions on cognitive tasks. Both the stepwise information gain and stepwise probability gain models are in good agreement with human subjects' eye movements on these tasks. In simulations, Najemnik and Geisler also found that if one's goal is to maximize classification accuracy (i.e., long-run probability gain), then under some circumstances, it is more efficient to use stepwise *information* gain than stepwise probability gain when choosing which location to view (fixate) in the next time step.

A related theoretical point is that even when a test is useless in a one-shot task, that does not mean that the test is useless as the first test in a multistep question strategy. For instance, it could be that the posterior probability of a disease is low enough (and the cost of treatment is high enough) that even if Test *E* is positive, it would not make sense to treat the patient; according to Pauker and Kassirer's influential (1980) analysis of medical diagnosis, such a test should not be conducted. But suppose that up to two tests can be conducted. Consider a

second test,  $F$ , that is also useless, according to the same analysis. It could be that if both tests are positive, the posterior probability of disease is above the threshold at which the patient should be treated. This example shows that a test can be useful if considered in conjunction with another test, even if both are useless on their own. A helpful experiment would investigate human test selection in a situation in which tests  $E$  and  $F$  are useless (relative to the objective utilities), considered on their own, and another test  $G$  is somewhat useful if considered on its own, but in which the combination of tests  $E$  and  $F$  is more useful than test  $G$ .

The above examples show the dangers of relying on stepwise strategies to quantify tests' usefulness. They also demonstrate surprising results wherein the best stepwise method can use an objective function that is different from the actual goal (see Singh, Lewis, Barto & Sorg, 2010, for investigation of a similar issue within the reinforcement-learning framework). The general issue of when and why particular stepwise strategies can “see the future” better than other strategies should be vigorously explored.

Such a theoretical investigation might also shed light on some strange findings in human test-selection behavior in the presence of asymmetric reward. There has not been much research in this area to date, but three papers have found that it is difficult for human subjects to choose tests that maximize utility, in the sense of overt reward, rather than pure information (like information gain or probability gain), on one-shot information acquisition tasks (Baron & Hershey, 1988; Markant & Gureckis, 2012; Meder & Nelson, 2012). Information about the structure of the world may be valuable in its own right (Filimon et al., 2016; Gottlieb et al., 2013), but information that is connected to reward (food, mating, predation, social status) is also important (Bromberg-Martin & Hikosaka, 2009).

It would be surprisingly maladaptive if people were unable to take situation-specific utilities into account when searching for information. One possibility is that people's apparently maladaptive behavior on these tasks could be an artifact of adaptation to finding

queries with high situation-specific utility in sequential search tasks. In other words, perhaps people conduct queries that, while not maximizing stepwise utility gain, somehow lead to high long run utility gain. At the moment, that is pure speculation, but it suggests avenues for future theoretical, behavioral, and cognitive neuroscience research. For instance, sequential search could be studied in real-world medical diagnosis tasks, to address whether apparently maladaptive preferences in one-shot tasks are adapted to sequential scenarios with external payoffs.

## **10.2 Need for empirical research on sequential planning of tests**

In statistics and computer science, consideration of stepwise approaches to question selection (e.g., Good, 1950; Lindley, 1956) came before consideration of efficiency in multistep question selection scenarios (e.g. Hyafil & Rivest, 1976). It is thus understandable that psychological research to date has focused on stepwise OED approaches. Future work on human information acquisition should explore the circumstances under which people can identify efficient long-run strategies. The body of relevant research to date is small: Meier and Blair (2013) found that people tended to first query a feature that led to higher efficiency, even though it was not the feature with highest stepwise probability gain or information gain; Meder et al. (2018) obtained much less promising results. What are the circumstances under which people can discover effective multistep strategies? As we highlight throughout this paper, there are a wealth of experiments that could be conducted to address this issue.

Although there is little research on sequential information search, there is research on whether and how people plan in other sequential decision tasks. For instance, Huys et al. (2012) studied a behavioral task in which people could transition between states, learning via experience the payoff or penalty associated with each state transition. Huys et al. modeled this in a reinforcement-learning framework (Sutton & Barto, 1998). Several models of human behavior were considered. One model used temporal discounting together with consideration of all possible future states. Such a model is feasible in small tasks, but faces a combinatorial

explosion in its computation when applied to larger tasks. Another model chose at random which states to prune and exclude from further consideration. A third model pruned only states that involved a large negative outcome; this model best accounted for human behavior.

These results could be used to help generate hypotheses for human behavior in sequential search tasks. For instance, people might not even consider Person Game question trees that begin with a highly unsplit halfy first question. Alternatively, the value of questions could be defined in terms of the quality of questions available at the next time step (e.g., as the expected split halfiness of the questions available at the next timestep), as in hybrid model-based–model-free approaches to reinforcement learning. As a research strategy, the goal could be to make these ideas computationally explicit, and in turn, to identify scenarios in which different models make strongly different predictions for behavioral experiments.

### **10.3 The quest for optimality: from rational models to heuristics and back**

In the literature on human information search, there has been little research on the ways in which the human cognitive system influences which algorithms are feasible. Historically, this may follow from Anderson's (1990) methodological suggestion to derive an optimal strategy solely from the task structure and the goals of the agent, making as few assumptions as possible about computational limitations. However, as we have illustrated in the case of sequential search, what is in fact optimal for human cognition is seldom clear (see also Jones & Love, 2011; Meder, Mayrhofer, & Waldmann, 2014). Heuristic strategies, as discussed in this manuscript, and boundedly rational models (Griffiths, Leider & Goodman, 2015), offer complementary avenues for developing predictive models.

What are the implications of our analyses and the juxtaposition of one-shot vs. sequential search scenarios? First, cognitive science should take seriously the fact that deriving the globally optimal solution is not computationally feasible for many problems. Importantly, as Person Game and Planet Vuma scenarios exemplify, this does not necessarily

result from limited knowledge of the structure of the environment. Even in situations in which all required information is available (hypotheses and features, the probabilistic structure of the environment, inter-feature dependencies, noiseless feedback, etc.), such as the full set of faces and features in a Person Game, the computational complexity can make it infeasible to calculate the globally optimal solution. For these problems, the question arises of what role the notion of optimality plays in providing guidance in model development and psychological theorizing. It is clear that some heuristic strategies are needed (see Hyafil & Rivest, 1976). But what kind of heuristic strategies? From a computer science standpoint, any algorithm that approximates a globally optimal algorithm yet has lower computational cost is a heuristic. In psychology there are varying definitions of, and views on, heuristics (e.g., Gigerenzer, 1996; Gigerenzer, Todd, & the ABC Research Group, 1999; Gilovich & Griffin, 2002; Kahneman & Tversky, 1974, 1996).

We suggest that which behavior should be considered optimal depends strongly on the cognitive and perceptual system of the agent. In the case of perception, this is readily taken into account. For instance, Butko and Movellan (2010) showed that depending on the foveal acuity curve (how rapidly peak resolution decreases as a function of distance from the center of gaze), one's eye movements to find a target object should be directed either to the most likely center of the target object, or slightly to the side of the target object. This implies that humans, who have a flat acuity curve within the fovea in the center of gaze (Duncan & Boynton, 2003; Sereno et al., 1995), should look slightly to the side of the most likely object location. However, an animal or computer vision system with a narrow point of peak acuity in the center of gaze should try to look at the center of the target object. We would not say that it is "biased" or "suboptimal" for eye movements to take the structure of the visual system into account. If beauty is in the eye of the beholder, then the rationality of an eye movement strategy is in the shape of the retinal acuity curve.

Some heuristics may not be optimal in the sense of always identifying the very most efficient strategies; rather, the heuristics may help people identify reasonable strategies more efficiently. Just as the properties of the visual system suggest which eye movement strategies are most efficient, properties of the cognitive system—for instance, the resources available in working memory, and how much a person intrinsically values his or her time—are relevant to cognitive psychological sequential search tasks. People’s use of information-search heuristics varies according to the information value of the heuristic and other task constraints (Coenen, Rehder & Gureckis, 2015; Klayman & Ha, 1989; Langsford, Hendrickson, Perfors & Navarro, 2014; Markant, Settles & Gureckis, 2015). Is a Person Game learner “suboptimal” or “biased” if he or she selects questions in a stepwise manner by using the split-half heuristic? Is it “optimal” for a learner to spend an extra 30 seconds, or an extra 30 minutes, to identify the highest two-step probability gain (or information gain) strategy? The answer depends on how much a person intrinsically values their time. Psychological models of question selection should take factors like the psychological cost of time into account in a meaningful quantitative way.

We suggest that OED models of information search, heuristics (and “biases”) reported in psychological literature, and the computer science literature on sequential search and other sequential planning tasks (e.g. using reinforcement learning) are equally valuable sources of inspiration for candidate descriptive cognitive models of human information acquisition. Whether a model was initially conceived as normative, adaptive, boundedly rational (Griffiths, Lieder, & Goodman, 2015), or even as a silly heuristic strategy, could almost be deemed irrelevant at this stage of research (also see Cohen, 1981; McKenzie, 1996; Tauber, Navarro, Perfors & Steyvers, 2017). From a methodological standpoint, for the goal of developing a descriptive model of human behavior, what is most important is that the models under consideration make precise predictions that can be differentiated from other models’ predictions via simulations, mathematical analyses, and behavioral experiments.

## Funding

This research was supported by grants ME 3717/2-2 to BM and NE 1713/1-2 to JDN from the Deutsche Forschungsgemeinschaft (DFG) as part of the priority program “New Frameworks of Rationality” (SPP 1516), by a University of Surrey Faculty Research Support Fund grant to JDN, and by and Air Force Office of Scientific Research grant FA9550-14-1-0318 to MJ.

## Acknowledgements

We thank Luisa Bentivegna, Bryan Bergert, Henry Brighton, Anna Coenen, Vincenzo Crupi, Flavia Filimon, Gerd Gigerenzer, Gudny Gudmundsdottir, Wasilios Hariskos, Laura Martignon, Javier Movellan (for discussion of Theorem 1), Felix Rebitschek, Azzurra Ruggeri, Özgür Şimşek, Katya Tentori, Jan Woike, Charley Wu, and the ABC Research Group for helpful contributions to this research and manuscript. We thank Andreas Sommer for designing the monster stimuli in Figures 1 and 6.

## References

- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Aström, K. J. (1965). Optimal control of Markov decision processes with incomplete state estimation. *Journal of Mathematical Analysis and Applications*, *10*, 174–205.
- Austerweil, J. L., & Griffiths, T. L. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, *35*, 499–526. doi: 10.1111/j.1551-6709.2010.01161.x
- Baron, J. (1985). *Rationality and intelligence*. Cambridge, England: Cambridge University Press.
- Baron, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: I. Priors, error costs, and test accuracy. *Organizational Behavior and Human Decision Processes*, *41*, 259–279. doi: 10.1016/0749-5978(88)90030-1

- Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, *42*, 88–110. doi: 10.1016/0749-5978(88)90021-0
- Bayes, T. (1763/1958). An essay towards solving a problem in the doctrine of chance. *Philosophical Transactions of the Royal Society of London*, *53*, 370-418. Reprinted in 1958 in *Biometrika*, *45*, 296-315.
- Bellman, R.E. 1957. *Dynamic Programming*. Princeton, NJ: Princeton University Press.
- Benish, W. A. (1999). Relative entropy as a measure of diagnostic information. *Medical Decision Making*, *19*, 202–206. doi: 10.1177/0272989X9901900211
- Berghman, L., Goossens, D., & Leus, R. (2009). Efficient solutions for Mastermind using genetic algorithms. *Computers & Operations Research*, *36*, 1880–1885. doi: 10.1016/j.cor.2008.06.004
- Bramley, N. R., Lagnado, D. A. & Speekenbrink, M. (2015). Forgetful conservative scholars - How people learn causal structure through interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 708–731. doi: 10.1037/xlm0000061
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. I. (1984). *Classification and regression trees*. Belmont, Calif.: Wadsworth.
- Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, *63*, 119–126. doi: 10.1016/j.neuron.2009.06.009
- Butko, N. J., & Movellan, J. R. (2010). Infomax Control of Eye Movements. *IEEE Transactions on Autonomous Mental Development*, *2*, 91–107. doi: 10.1109/TAMD.2010.2051029
- Chakaravarthy, V. T., Pandit, V., Roy, S., Awasthi, P., & Mohania, M. (2007). Decision trees for entity identification: Approximation algorithms and hardness results. In *Proceedings*



- of the ACM-SIGMOD Symposium on Principles of Database Systems* (pp. 53-62). ACM Digital Library.
- 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*, 63–86. doi: 10.1016/S0749-5978(02)00508-3
- Coenen, A., Nelson, J. D., & Gureckis, T. M. (2018). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin and Review*. doi: 10.3758/s13423-018-1470-5
- Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology*, *79*, 102–133. doi: 10.1016/j.cogpsych.2015.02.004
- Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory*. Wiley, New York, NY.
- Crupi, V., & Tentori, K. (2014). State of the field: Measuring information and confirmation. *Studies in History and Philosophy of Science*, *47(C)*, 81–90. doi: 10.1016/j.shpsa.2014.05.002
- Crupi, V., Nelson, J. D., Meder, B., Cevolani G., & Tentori, K. (2018). Generalized information theory meets human cognition: Introducing a new framework to model uncertainty and information search. *Cognitive Science*, *42*, 1410-1456.
- Crupi, V., Tentori, K., & Lombardi, L. (2009). Pseudodiagnosticity revisited. *Psychological Review*, *116*, 971–985. doi: 10.1037/a0017050
- Daw, N.D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*, 1704-1711.

- Denney, D. R., & Denney, N. W. (1973). The use of classification for problem solving: A comparison of middle and old age. *Developmental Psychology, 9*, 275–278. doi: 10.1037/h0035092
- Denzler, J., & Brown, C. M. (2002). Information theoretic sensor data selection for active object recognition and state estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*, 1–13. doi: 10.1109/34.982896
- Domingos, P. & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning, 29*, 103–130. doi: 10.1023/A:1007413511361
- Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron, 38*, 659–671. doi: 10.1016/S0896-6273(03)00265-4
- Eimas, P. D. (1970). Information processing in problem solving as a function of developmental level and stimulus saliency. *Developmental Psychology, 2*, 224–229. doi: 10.1037/h0028746
- Filimon, F., Nelson, J. D., Sejnowski, T. J., Sereno, M. I., & Cottrell, G. W. (2016). Dissociation between information expectation, reward anticipation, and reward receipt in the ventral striatum. Manuscript submitted for publication.
- Garey, M. R. (1972). Optimal binary identification procedures. *SIAM Journal on Applied Mathematics, 23*, 173–186. doi: 10.1137/0123019
- Geman, D., & Jedynek, B. (2001). Model-based classification trees. *IEEE Transactions on Information Theory, 47*, 1075–1082. doi: 10.1109/18.915664
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review, 103*, 592–596. doi: 10.1037/0033-295X.103.3.592
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review, 102*, 684–704.

- Gigerenzer, G., & Todd, P. M. the ABC Research Group 1999. *Simple heuristics that make us smart*. Oxford University Press.
- Gilovich, T., & Griffin, D. (2002). Introduction-- heuristics and biases, then and now. In Gilovich, T., Griffin, D., & Kahneman, D. (Eds.), *Heuristics and Biases: The Psychology of Intuitive Judgment*, (pp. 1-18). Cambridge, UK: Cambridge University Press.
- Golovin, D., & Krause, A. (2011). Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42, 427-486.
- Golovin, D., Krause, A., & Ray, D. (2010). Near-optimal Bayesian active learning with noisy observations. In Lafferty, J. D., Williams, C. K. I., Shawe-Taylor, J., Zemel, R. S., & Culotta, A. (Eds.), *Advances in Neural Information Processing Systems*, 23, 766-774.
- Good, I. J. (1950). *Probability and the weighing of evidence*. New York: Griffin.
- Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in Cognitive Sciences*, 17, 585-593. doi: 10.1016/j.tics.2013.09.001
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7, 217-229. doi: 10.1111/tops.12142
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7, 217-229.
- Gureckis, T. M. & Love, B. C. (2009). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition*, 113, 293-313. doi: 10.1016/j.cognition.2009.03.013

- Gureckis, T. M., & Markant, D. B. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, 7, 464–481. doi: 10.1177/1745691612454304
- Hotaling, J. M., Busemeyer, J. R. (2012). DFT-D: a cognitive-dynamical model of dynamic decision making. *Synthese*, 189(1), 67-80.
- Howard, R. A. (1960). *Dynamic programming and Markov processes*. Cambridge, MA: The M.I.T. Press.
- Huffman, D. A. (1952). A method for the construction of minimum redundancy codes. *Proceedings of the I.R.E.*, 40, 1098–1101. doi: 10.1109/JRPROC.1952.273898
- Huys, Q. J. M., Eshel, N., O’Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, 8(3): e1002410. doi:10.1371/journal.pcbi.1002410
- Hyafil, L., & Rivest, R. L. (1976). Constructing optimal binary decision trees is NP-complete. *Information Processing Letters*, 5, 15–17. doi: 10.1016/0020-0190(76)90095-8
- Jarecki, J. B., Meder, B., & Nelson, J. D. (2018). Naïve and robust: Class-conditional independence in human classification learning. *Cognitive Science*, 42, 4-42. doi: 10.1111/cogs.12496
- Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, 34, 169–188. doi: 10.1017/S0140525X10003134
- Kachergis, G., Rhodes, M., & Gureckis, T. (2016). Desirable difficulties in the development of active inquiry skills. *Cognition*, 166, 407-417. doi: 10.1016/j.cognition.2017.05.021
- Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101, 99–134.

- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, *103*, 582–591. doi: 10.1037/0033-295X.103.3.582
- Klayman, J. (1995). Varieties of confirmation bias. In J. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Psychology of Learning and Motivation*, *32*, 385–418. doi: 10.1016/S0079-7421(08)60315-1
- Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228. doi: 10.1037/0033-295X.94.2.211
- Klayman, J., & Ha, Y.-W. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 596–604. doi: 0.1037/0278-7393.15.4.596
- Knuth, D. E. (1971). Optimum binary search trees. *Acta Informatica*, *1*, 14–25. doi:10.1007/BF00264289
- Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, *45*, 812–863. doi:10.1006/jmps.2000.1354
- Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*, 210–226. doi: 10.3758/LB.36.3.210
- Langsford, S., Hendrickson, A., Perfors, A., & Navarro, D. J. (2014). People are sensitive to hypothesis sparsity during category discrimination. In P. Bellow, M. Guarani, M. McShane, and B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 2531–2536). Austin, Texas: Cognitive Science Society.
- Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: an ideal-observer model of reading. *Psychological Review*, *104*, 524–553. doi: 10.1037/0033-295X.104.3.524
- Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, *27*, 986–1005. doi: 10.1214/aoms/1177728069

- Markant, D. B., Settles, B., & Gureckis, T. M. (2015). Self-directed learning favors local, rather than global, uncertainty. *Cognitive Science*. doi: 10.1111/cogs.12220
- Markant, D., & Gureckis, T. M. (2012). Does the utility of information influence sampling behavior? In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 719–724). Austin, TX: Cognitive Science Society.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: Freeman.
- Mastropasqua, T., Crupi, V., & Tentori, K. (2010). Broadening the study of inductive reasoning: Confirmation judgments with uncertain evidence. *Memory & Cognition*, 38, 941–950. doi: 10.3758/MC.38.7.941
- McKenzie, C. R. M. (2006). Rational models as theories, not standards, of behavior. *Trends in Cognitive Sciences*, 7, 403–406. doi: 10.1016/S1364-6613(03)00196-7
- Meder, B., & Nelson, J. D. (2012). Information search with situation-specific reward functions. *Judgment and Decision Making*, 7, 119–148.
- Meder, B., Gerstenberg, T., Hagmayer, Y., & Waldmann, M. R. (2010). Observing and intervening: Rational and heuristic models of causal decision making. *The Open Psychology Journal*, 3, 119–135.
- Meder, B., Mayrhofer, R., & Waldmann, M. R. (2014). Structure induction in diagnostic causal reasoning. *Psychological Review*, 121, 277–301. doi: 10.1037/a0035944
- Meder, B., Nelson, J. D., Jones, M., Ruggeri, A. (2018). Stepwise versus globally optimal search in children and adults. Manuscript submitted for publication.
- Meier, K. M., & Blair, M. R. (2013). Waiting and weighting: Information sampling is a balance between efficiency and error-reduction. *Cognition*, 126, 319–325. doi: 10.1016/j.cognition.2012.09.014

- Mitchell, T. M. (1982). Generalization as search. *Artificial Intelligence*, 18(2), 203–226.  
doi:10.1016/0004-3702(82)90040-6
- Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Computational Biology*, 8, e1002342.  
doi: 10.1371/journal.pcbi.1002342
- Mosher, F. A., & Hornsby, J. R. (1966). On asking questions. In J. S. Bruner, R. R. Olver, P., and M. Greenfield (Eds.), *Studies in cognitive growth* (pp. 86–102). New York, NY: Wiley.
- Myung, J. I., Cavagnaro, D. R., & Pitt, M. A. (2013). A tutorial on adaptive design optimization. *Journal of Mathematical Psychology*, 57, 53–67.  
doi: 10.1016/j.jmp.2013.05.005
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434, 387–391. doi: 10.1038/nature03390
- Nakamura, K. (2006). Neural representation of information measure in the primate premotor cortex. *Journal of Neurophysiology*, 96, 478–485. doi: 10.1152/jn.01326.2005
- Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review*, 118, 120–134. doi: 10.1037/a0021110
- Navarro, D. J., Newell, B. R., & Schulze, C. (2016). Learning and choosing in an uncertain world: An investigation of the explore–exploit dilemma in static and dynamic environments. *Cognitive Psychology*, 85, 43–77. doi: 10.1016/j.cogpsych.2016.01.001
- Nelson, J. D. (2005a). Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, 112, 979–999. doi: 10.1037/0033-295X.112.4.979
- Nelson, J. D. (2005b). *Optimal Experimental Design as a Theory of Perceptual and Cognitive Information Acquisition*. PhD dissertation, Cognitive Science Department, University of California, at San Diego.

- Nelson, J. D., Divjak, B., Gudmundsdottir, G., Martignon, L. F., & Meder, B. (2014). Children's sequential information search is sensitive to environmental probabilities. *Cognition, 130*, 74–80. doi: 10.1016/j.cognition.2013.09.007
- Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience matters: Information acquisition optimizes probability gain. *Psychological Science, 21*, 960–969. doi: 10.1177/0956797610372637
- Nelson, J. D., Tenenbaum, J. B., & Movellan, J. R. (2001). Active inference in concept learning. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Conference of the Cognitive Science Society* (pp. 692–697). Mahwah, NJ: Erlbaum.
- Nemhauser, G. L., Wolsey, L. A., & Fisher, M. L. (1978). An analysis of approximations for maximizing submodular set functions - I. *Mathematical Programming, 14*(1), 265–294.
- Newell, A., & Simon, H. A. (1959). *The simulation of human thought*. Santa Monica, CA: Rand Corp.
- Nickerson, R. S. (1996). Hempel's paradox and Wason's selection task: Logical and psychological puzzles of confirmation. *Thinking & Reasoning, 2*, 1–31. doi: 10.1080/135467896394546
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608–631. doi: 10.1037/0033-295X.101.4.608
- Oaksford, M., & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review, 103*, 381–391. doi: 10.1037/0033-295X.103.2.381
- Oaksford, M., & Chater, N. (2003). Optimal data selection: Revision, review, and reevaluation. *Psychonomic Bulletin & Review, 10*, 289–318. doi: 10.3758/BF03196492
- Pauker, S. G., & Kassirer, J. P. (1980). The threshold approach to clinical decision making. *The New England Journal of Medicine, 302*, 1109–1117. doi: 10.1056/NEJM198005153022003.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.



- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*, 81–106.  
doi: 10.1023/A:1022643204877
- Raiffa, H., & Schlaifer, R. O. (1961). *Applied Statistical Decision Theory*. Cambridge, MA: Division of Research, Graduate School of Business Administration, Harvard University.
- Renninger, L. W., Coughlan, J., Verghese, P., & Malik, J. (2005). An information maximization model of eye movements. *Advances in Neural Information Processing Systems, 17*, 1121–1128.
- Ruggeri, A., & Lombrozo, T. (2015). Children adapt their questions to achieve efficient search. *Cognition, 143*, 203-216.
- Ruggeri, A., Lombrozo, T., Griffiths, T. L., & Xu, F. (2016). Sources of developmental change in the efficiency of information search. *Developmental Psychology, 52*, 2159-2173.
- Ruggeri, A., Sim, Z. L., & Xu, F. (2017). “Why is Toma late to school again?” Preschoolers identify the most informative questions. *Developmental Psychology, 53*, 1620-1632.  
doi: 10.1037/dev0000340
- Rusconi, P., Marelli, M., D’Addario, M., Russo, S., & Cherubini, P. (2014). Evidence evaluation: Measure Z corresponds to human utility judgments better than measure L and optimal-experimental-design models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 703–723. doi: 10.1037/a0035549
- Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*, 889–893. doi: 10.1126/science.7754376
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*, 379–423, 623–656.

- Siegler, R. S. (1977). The twenty questions game as a form of problem solving. *Child Development*, 48, 395–403. doi: 10.2307/1128632
- Singh, S., Lewis, R. L., Barto, A. G., & Sorg, J. (2010). Intrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE Transactions on Autonomous Mental Development*, 2, 70–82. doi: 10.1109/TAMD.2010.2051031
- Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, 103, 278–282. doi:10.1016/j.beproc.2014.01.014
- Slowiaczek, L. M., Klayman, J., Sherman, S. J., & Skov, R. B. (1992). Information selection and use in hypothesis testing: What is a good question, and what is a good answer? *Memory & Cognition*, 20, 392–405. doi:10.3758/BF03210923
- Steyvers, M., Tenenbaum, J. B., Wagenmakers, E. J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, 27, 453–489. doi: 10.1016/S0364-0213(03)00010-7
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Tauber, S., Navarro, D. J., Perfors, A., & Steyvers, M. (2017). Bayesian models of cognition revisited: Setting optimality aside and letting data drive psychological theory. *Psychological Review*, 124, 410–441. <http://dx.doi.org/10.1037/rev0000052>
- Thornton, S. (1982). Challenging “early competence”: A process oriented analysis of children's classifying. *Cognitive Science*, 6, 77–100. doi: 10.1207/s15516709cog0601\_3
- Todd, P. M., & Gigerenzer, G., & the ABC Research Group. (2012). *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.
- Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology*, 43, 22–34. doi: 10.1037/0022-3514.43.1.22

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131, doi: 10.1126/science.185.4157.1124.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*, 129–140. doi: 10.1080/17470216008416717
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New Horizons in Psychology* (pp. 135–151). Harmondsworth, England: Penguin.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273–281. doi: 10.1080/14640746808400161
- Wells, G. L., & Lindsay, R. C. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, *88*, 776–784. doi: 10.1037/0033-2909.88.3.776
- Wu, C. M., Meder, B., Filimon, F., & Nelson, J. D. (2017). Asking better questions: How presentation formats influence information search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1274-1297. doi: 10.1037/xlm0000374
- Wu, C., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behavior*. doi: 10.1038/s41562-018-0467-4
- Zhang, S., & Yu, A. J. (2013). Forgetful Bayes and myopic planning: Human learning and decision-making in a bandit setting. *Advances in Neural Information Processing Systems*, *26*, 2607-2615.

### Tables

Table 1

*Illustration of a Person Game Environment, with Eight Hypotheses (Persons),  $h_1, \dots, h_8$ , and Seven Binary Features that can be Queried,  $Q_1, \dots, Q_7$ , as well as the Stepwise Information Values of Each First Question*

Person	$Q_1$	$Q_2$	$Q_3$	$Q_4$	$Q_5$	$Q_6$	$Q_7$
$h_1$	1	0	1	0	0	0	1
$h_2$	1	0	1	0	0	0	0
$h_3$	1	0	1	0	0	1	0
$h_4$	1	0	0	0	0	0	0
$h_5$	0	0	0	0	0	0	0
$h_6$	0	1	0	0	0	0	0
$h_7$	0	1	0	0	1	0	0
$h_8$	0	1	0	1	0	0	0
Utility function							
Probability gain	.125	.125	.125	.125	.125	.125	.125
Information gain	1.000	.954	.954	.544	.544	.544	.544
Impact	1.000	.938	.938	.438	.438	.438	.438

*Note.* Figure 1 gives example stimuli instantiating this structure. Figure 3 shows different question trees that might be used on this task. In the Person Game, all name questions are allowed (e.g., “Is it  $h_1$ ?”). We do not explicitly include the name questions for  $h_2, h_4, h_5$ , or  $h_6$  in this table, because they are not needed for constructing the example trees in Figure 3. For the expected value of querying each feature as a first question, within a row (a particular OED model), larger values correspond to higher usefulness. Information gain is measured in bits.

Table 2

*Information Value of Four Selected Person Game Trees (from Figure 3, Based on the Environment in Table 1), According to Several Multistep OED Models*

	Stepwise-optimal tree	Globally optimal tree	Start with name question tree	Name questions only tree
Expected path length				
Full tree	3.250	3.125	3.625	4.375
Expected impact				
1-step	1.000	0.938	0.438	0.438
2-step	1.375	1.438	1.188	0.813
3-step	1.625	1.688	1.500	1.125
Full tree	1.750	1.750	1.750	1.750
Expected information gain				
1-step	1.000	0.954	0.544	0.544
2-step	1.811	1.906	1.406	1.061
3-step	2.500	2.750	2.156	1.545
Full tree	3.000	3.000	3.000	3.000
Expected probability gain				
1-step	0.125	0.125	0.125	0.125
2-step	0.375	0.375	0.250	0.250
3-step	0.625	0.750	0.500	0.375
Full tree	0.875	0.875	0.875	0.875

*Note.* Figure 3, top row, depicts the full trees. Figure 3, middle row, depicts the subtrees that are defined by the first 2 steps of the corresponding full trees. Figure 3, bottom row, depicts the subtrees that are defined by the first 3 steps of the corresponding full trees. Path length = number of queries to reach certainty. Information gain = expected reduction in (Shannon) entropy, in bits. Probability gain = expected improvement in classification accuracy. Impact = absolute change in beliefs.

Table 3

*Constraint Questions are not Always Better than Hypothesis-Scanning Questions*

	Constraint questions			Name questions									
	$Q_1$	$Q_2$	$Q_3$	$Q_4$	$Q_5$	$Q_6$	$Q_7$	$Q_8$	$Q_9$	$Q_{10}$	$Q_{11}$	$Q_{12}$	$Q_{13}$
$h_1$	0	0	0	1	0	0	0	0	0	0	0	0	0
$h_2$	0	0	0	0	1	0	0	0	0	0	0	0	0
$h_3$	0	0	0	0	0	1	0	0	0	0	0	0	0
$h_4$	0	0	0	0	0	0	1	0	0	0	0	0	0
$h_5$	0	0	0	0	0	0	0	1	0	0	0	0	0
$h_6$	0	0	0	0	0	0	0	0	1	0	0	0	0
$h_7$	0	1	1	0	0	0	0	0	0	1	0	0	0
$h_8$	1	0	0	0	0	0	0	0	0	0	1	0	0
$h_9$	0	1	0	0	0	0	0	0	0	0	0	1	0
$h_{10}$	1	0	1	0	0	0	0	0	0	0	0	0	1
Mean path length	4.2	4.2	4.7	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.8	4.8	4.8

*Note.* Figure 6 instantiates this environment graphically. In the developmental psychology literature, constraint questions are deemed superior to hypothesis-scanning questions. This example environment shows that not even that basic ordering holds in general. The bottom row shows the expected number of questions required to identify the target if the corresponding question is asked first, provided that subsequent questions are chosen according to the optimal strategy. The remaining rows indicate which hypotheses (persons) have each feature. Most hypothesis-scanning questions, namely  $Q_4$  through  $Q_9$ , are superior to the constraint question  $Q_3$ .

Table 4

*Trees of Depth Two in the Two-Category Binary-Feature (Planet Vuma) Example*

Tree	First test	Second test if first test $e_1$ or $a_1$	Second test if first test $e_2$ or $a_2$	Tree path length	Tree prob. gain	Tree info. gain	Tree impact
Globally Optimal	E	--	E	1.601	.171	0.502	0.685
Stepwise Probability Gain	A	E	A	2.000	.146	0.274	0.470
Other trees:	A	A	A	2.000	.130	0.241	0.451
	E	--	A	1.601	.102	0.365	0.479
	A	E	E	2.000	.102	0.365	0.479
	A	A	E	2.000	.086	0.332	0.460

*Note.* This example shows how stepwise procedures for choosing queries can be suboptimal, even in situations with just two hypotheses. Each of the six possible trees is described, together with its expected path length and its value according to (expected) probability gain, information gain, and impact. The globally optimal tree (Figure 8d) is optimal with respect to all three two-step OED criteria and also ties for requiring the fewest questions. Applying information gain, impact, or the likelihood difference heuristic in a stepwise manner all result in this tree. The stepwise probability gain tree (Figure 8c) results from applying probability gain in a stepwise manner. The other four possible trees are also included for comparison purposes. The trees are ordered from highest to lowest probability gain, with the shorter path length tree listed first in the case of a tie.

Figures

Figure 1. Stimuli to illustrate the Person Game environment in Table 1. The goal is to use the strategy that identifies the correct monster with the fewest expected queries. Allowable questions are about the physical features of those monsters, including the features that individuate single monsters, such as “Does the monster have a bow tie?”. The shape question gives a perfect 4:4 split. Although the color question and the eye question each give only a 3:5 initial split, they are superior questions from the perspective of long-run efficiency.

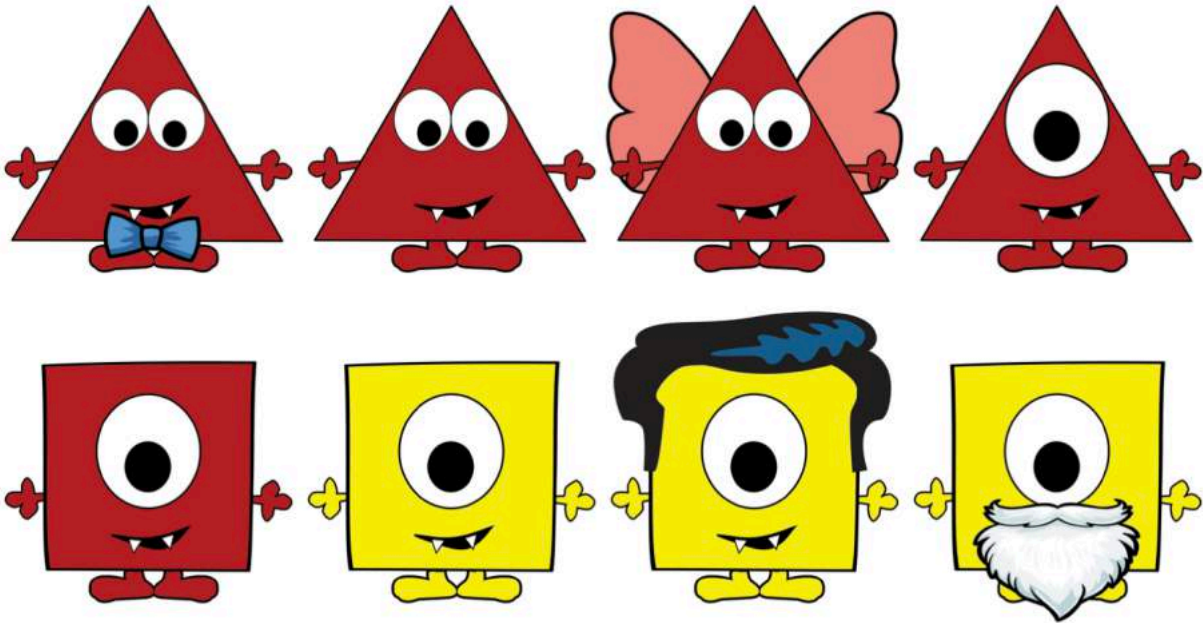




Figure 2. Usefulness of questions in the Person Game, according to stepwise probability gain, information gain and impact, assuming a uniform prior over the hypotheses. The curves are formed by taking possible probabilities of a feature being present,  $P(q_{\text{yes}})$ , and calculating the usefulness of asking about that feature, according to each OED model. Each curve is scaled to a maximum of 1. The information gain and impact curves are symmetric, with maxima at  $P(q_{\text{yes}}) = .5$ , implying that the split-half heuristic can be used to identify the highest impact and highest information gain questions, irrespective of the number of items being guessed among. The probability gain curve is flat and maximal for all values of  $P(q_{\text{yes}})$ ,  $0 < P(q_{\text{yes}}) < 1$ , and zero if  $P(q_{\text{yes}}) = 0$  or  $P(q_{\text{yes}}) = 1$ .

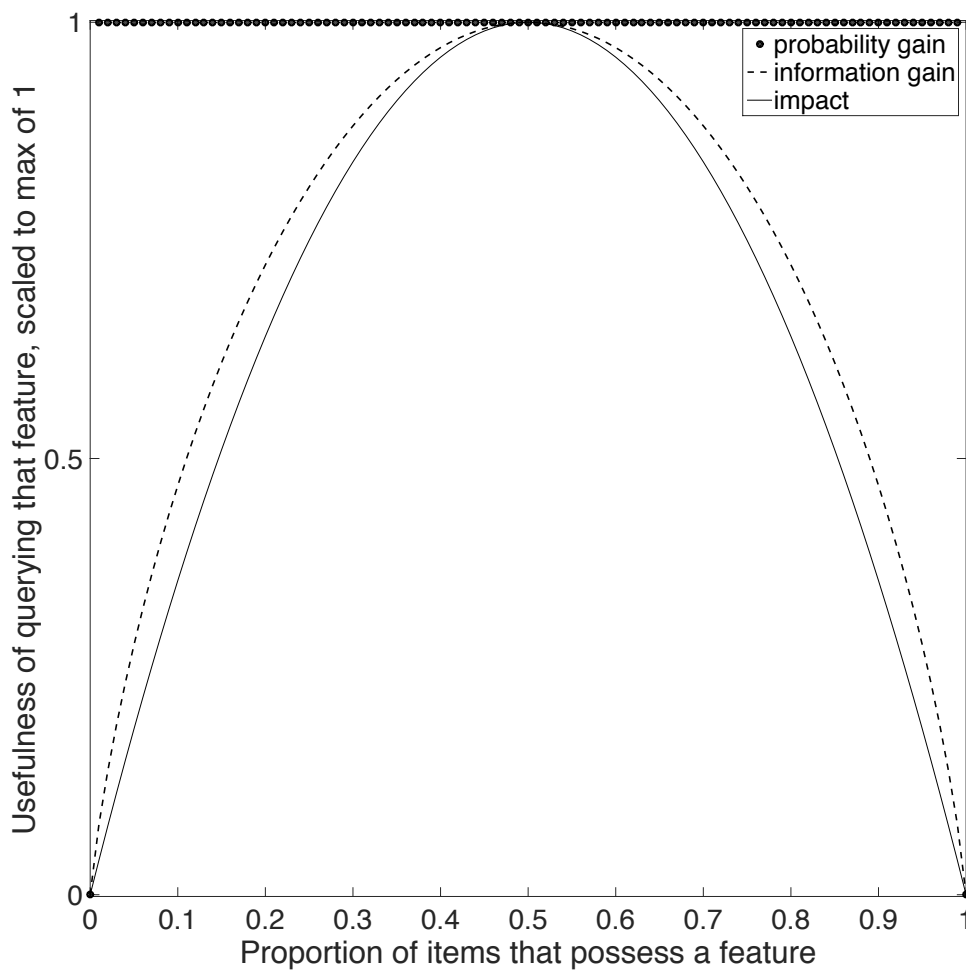
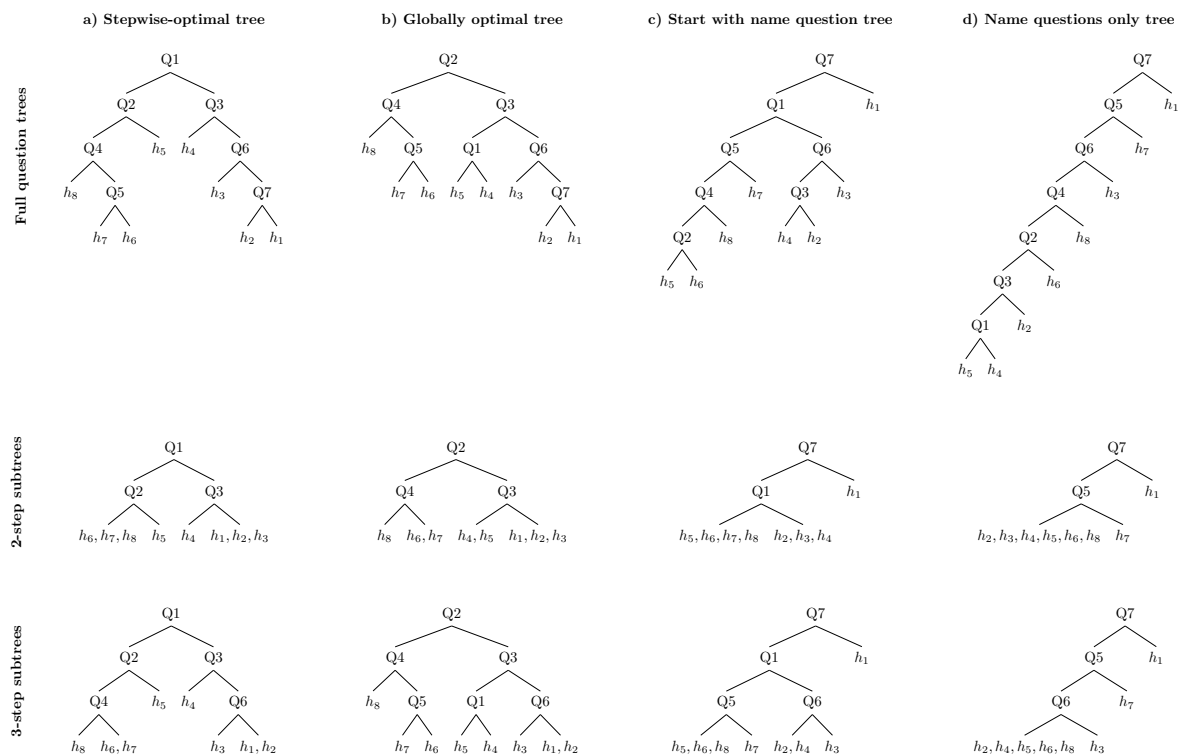
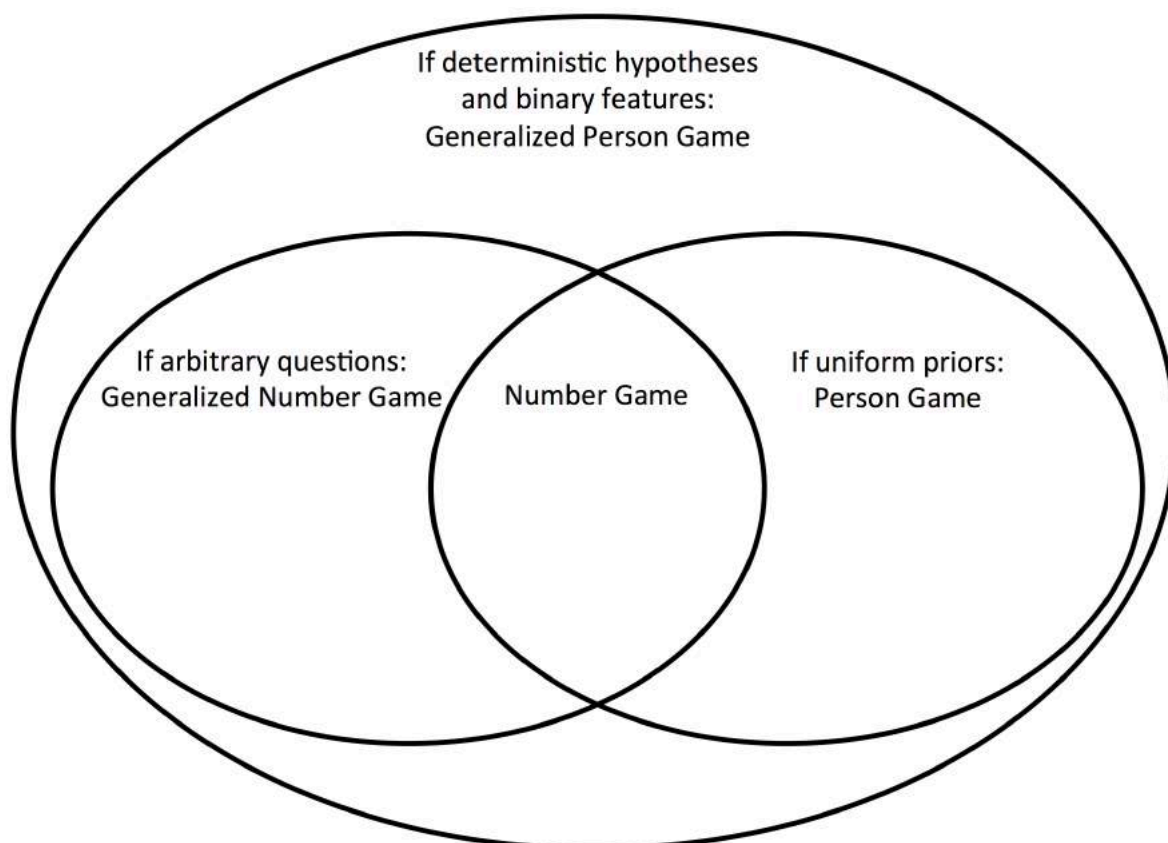


Figure 3. Four possible question strategies (trees) for the Person Game environment in Table 1, which is instantiated graphically in Figure 1. If the answer to a query is negative, the left branch is taken; in case of a positive answer the right branch is taken. The “stepwise optimal” tree (a) results from applying the stepwise methods (split-half heuristic, maximal stepwise expected impact, or maximal stepwise expected information gain). The “globally optimal” tree (b) achieves certainty in the smallest expected number of questions. (Note that two trees—starting with either Q2 or Q3—tie for globally optimal, so we arbitrarily depict one of them here.) The “start with name question” tree (c) and the “name questions only” tree (d) illustrate different possible trees that start by asking a hypothesis-scanning or “name” question (querying a feature unique to one hypothesis). The middle row depicts the subtrees of depth 2, used for evaluating 2-step OED models; the bottom row depicts the subtrees of depth 3, for evaluating 3-step OED models. Table 2 gives results on the efficiency of each tree, as well as the expected impact, information gain, and probability gain that can be achieved from the first one, two, or three steps of each tree.



*Figure 4.* Illustration of how different kinds of Generalized Person Game relate to each other. In each type of game, the task is to identify an individual person (or item). Each item is uniquely characterized by a set of binary attributes. Each variant of the game is characterized by what kind of questions are available to ask, and by the prior probability distribution of the items. Every version of the game is a Generalized Person Game. If the prior probabilities of all items are equal, then the game is a Person Game. If every possible question can be asked (e.g., if the questioner can ask whether the target item is part of any arbitrary subset), then the game is a Generalized Number Game (which we do not specifically discuss). If all of the above conditions hold, then the game is a Number Game. Theorems 1 and 8 apply to the Generalized Person Game, and therefore also the Person Game, Generalized Number Game, and Number Game.



*Figure 5.* Highlights of results from Person Game simulations. (a) Probability that the stepwise methods (split-half heuristic, maximal stepwise impact, and maximal stepwise information gain) would select a suboptimal first question as the first question of various Person Game environments. Results are based on taking the mean value of the % Suboptimal Choices column in Figures C1 through C4, for each size Person Game. Performance of the stepwise methods is the worst when features tend to be correlated with each other and are diverse in their densities. (b) Probability that the stepwise methods select a suboptimal first question as a function of number of constraint questions (averaged over environment sizes). Performance is worst in environments with a small number of constraint features. If there were 20 constraint features, it was rare for the stepwise methods to select a suboptimal first query (within the range of environment sizes tested). Figures C1 through C4 give a more detailed breakdown of the performance of the stepwise methods as a function of the number of constraint questions and average feature density, for each size Person Game. Features were generated by four methods: reg. features = regular features (independent and identically distributed); div. features = diverse features (independent with varying densities); corr. features = correlated features (with equal density); corr. div. features = correlated diverse features.

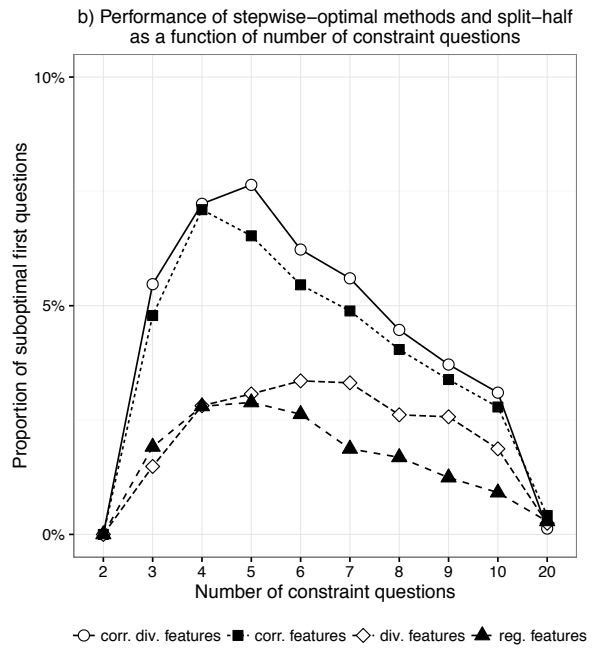
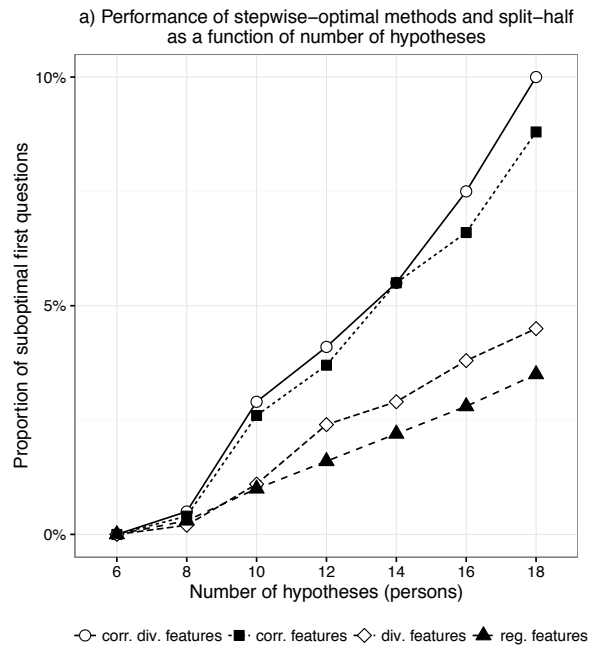


Figure 6. Experimental stimuli to illustrate that constraint questions are not always better than name questions, instantiating the Person Game environment in Table 3.

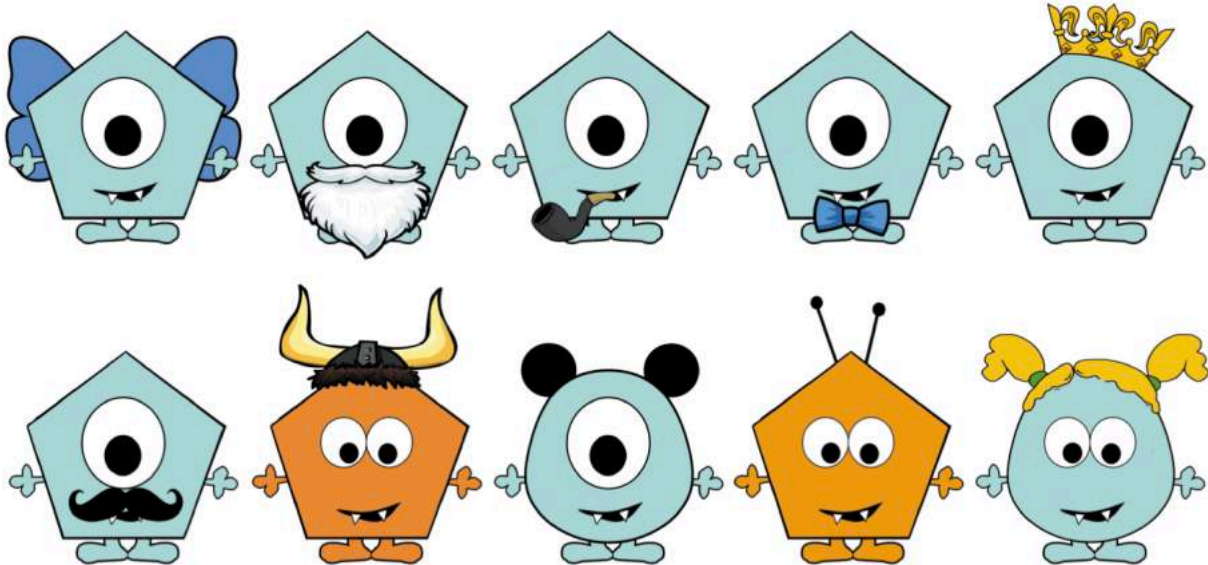
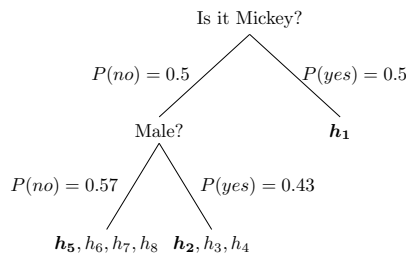
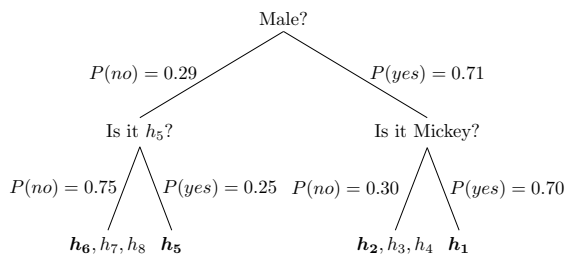


Figure 7. Two-step trees for the Mickey Mouse example Generalized Person Game, in which prior probability is  $\frac{1}{2}$  for one hypothesis ( $h_1$ : Mickey) and  $1/14$  for the others. The split-half trees yield greater two-step probability gain than does the stepwise information gain tree. In each tree, the player's best guess at each leaf node, meaning the hypothesis with greatest probability, is shown in boldface (chosen arbitrarily in cases of ties). These highlighted hypotheses correspond to the  $h_j^*$  defined in Theorem 8. According to that theorem, the expected accuracy afforded by each tree equals the sum of prior probabilities of these best guesses. Thus the stepwise information gain tree has accuracy  $p(h_1) + p(h_2) + p(h_5) = 9/14$ , whereas both split-half trees have accuracy  $p(h_1) + p(h_2) + p(h_5) + p(h_6) = 10/14$ . Because initial accuracy is  $1/2$  (guessing Mickey without any queries), the respective probability gains are  $2/14$  and  $3/14$ .

a) Maximum entropy and stepwise information gain tree



b) Split-half tree I



c) Split-half tree II

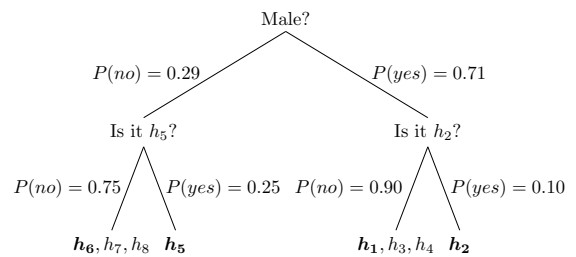


Figure 8. Stepwise versus two-step approaches in a two-category information-acquisition paradigm. (a) Statistical structure of the example environment:  $P(h_1) = .70$  and  $P(h_2) = .30$ , and likelihoods  $P(a_1 | h_1) = .04$ ,  $P(a_1 | h_2) = .38$ ,  $P(e_1 | h_1) = .57$ , and  $P(e_1 | h_2) = 0$ . The  $A$  and  $E$  tests are class-conditionally independent (see text for full description and intuitive explanation). (b) In the typical (“Planet Vuma”) task, just a single test (Test  $A$  or Test  $E$ ) can be conducted. Test  $A$  leads to higher accuracy (and hence probability gain) for this environment, but Test  $E$  has higher information gain and impact. (c) If two tests can be conducted, stepwise probability gain first conducts Test  $A$ . If the result is  $a_1$ , the second test is Test  $E$ ; if the result is  $a_2$ , Test  $A$  is used a second time. (d) If two tests can be conducted, the optimal tree starts with Test  $E$ . If the first test’s result is  $e_1$ , the category is known to be  $h_1$ ; if the first test’s result is  $e_2$ , Test  $E$  is repeated. This tree has highest overall accuracy, information gain, and impact. Moreover, among all six trees with up to two tests (Table 4), this tree ties for requiring the fewest queries, on average. Stepwise information gain, stepwise impact, and the likelihood difference heuristic each follow this tree.

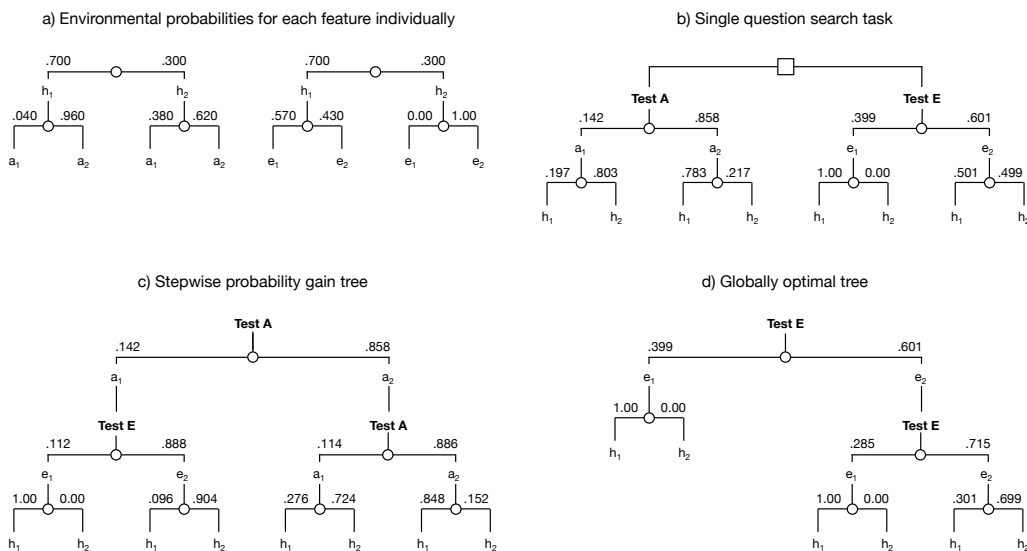




Figure C1. Detailed simulation results for many Person Game environment simulations, with regular features. “Regular features” means that features are generated independently of each other, all with the same generating density. The left column is a baseline and gives the number of questions required by the optimal strategy. The next column gives the percentage of instances in which the stepwise methods select a suboptimal first question. The right column gives the average cost, in expected number of questions, of using the stepwise methods (averaged over cases in which it is optimal and in which it is not). Each pixel represents the mean across 100 different random Person Game environments.

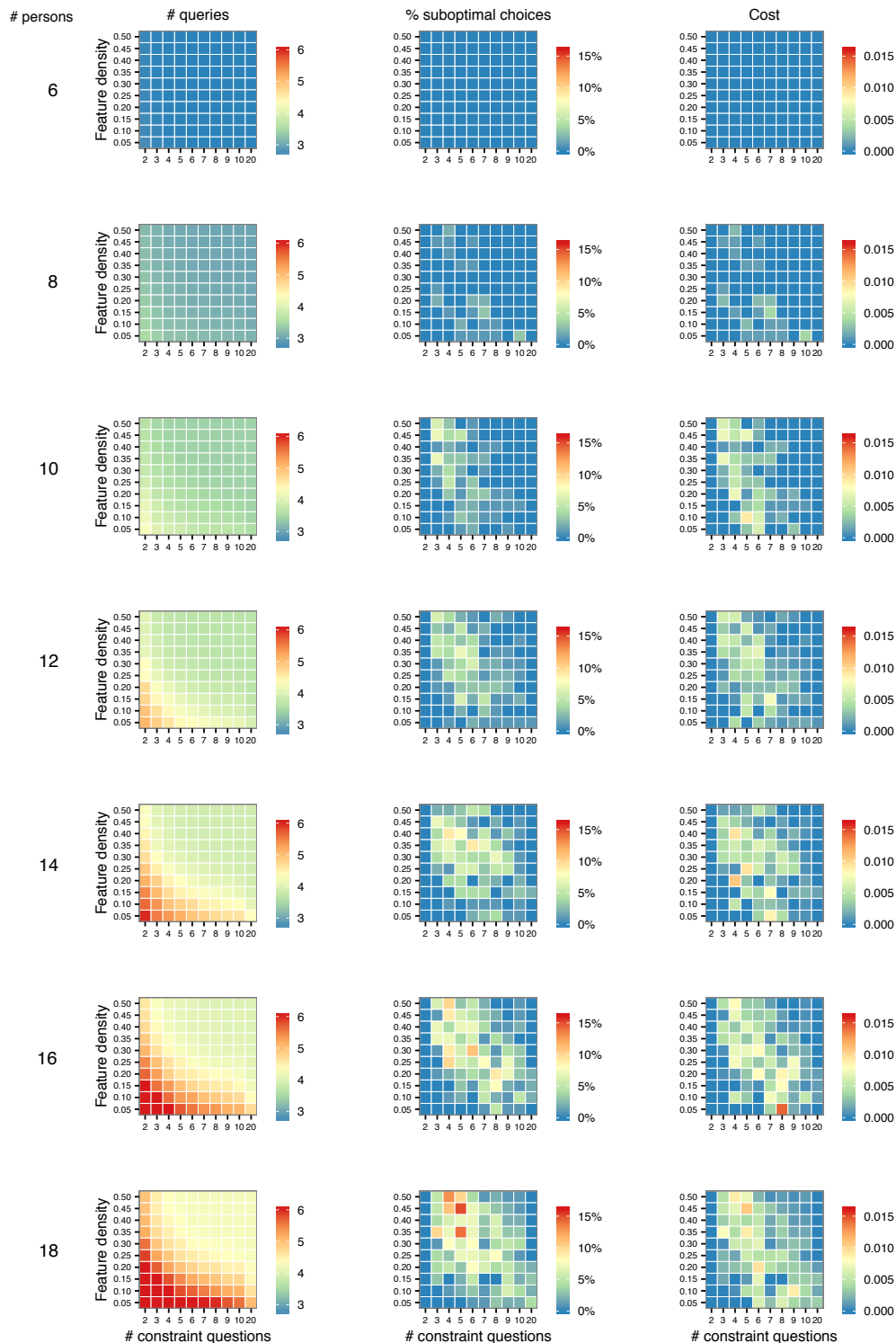


Figure C2. Detailed simulation results for many Person Game environment simulations, with diverse features. “Diverse features” means that each feature has a generating feature density that is sampled independently of the other features, from a Beta distribution with specified mean feature density. Other details of the plot are the same as explained in Figure C1.

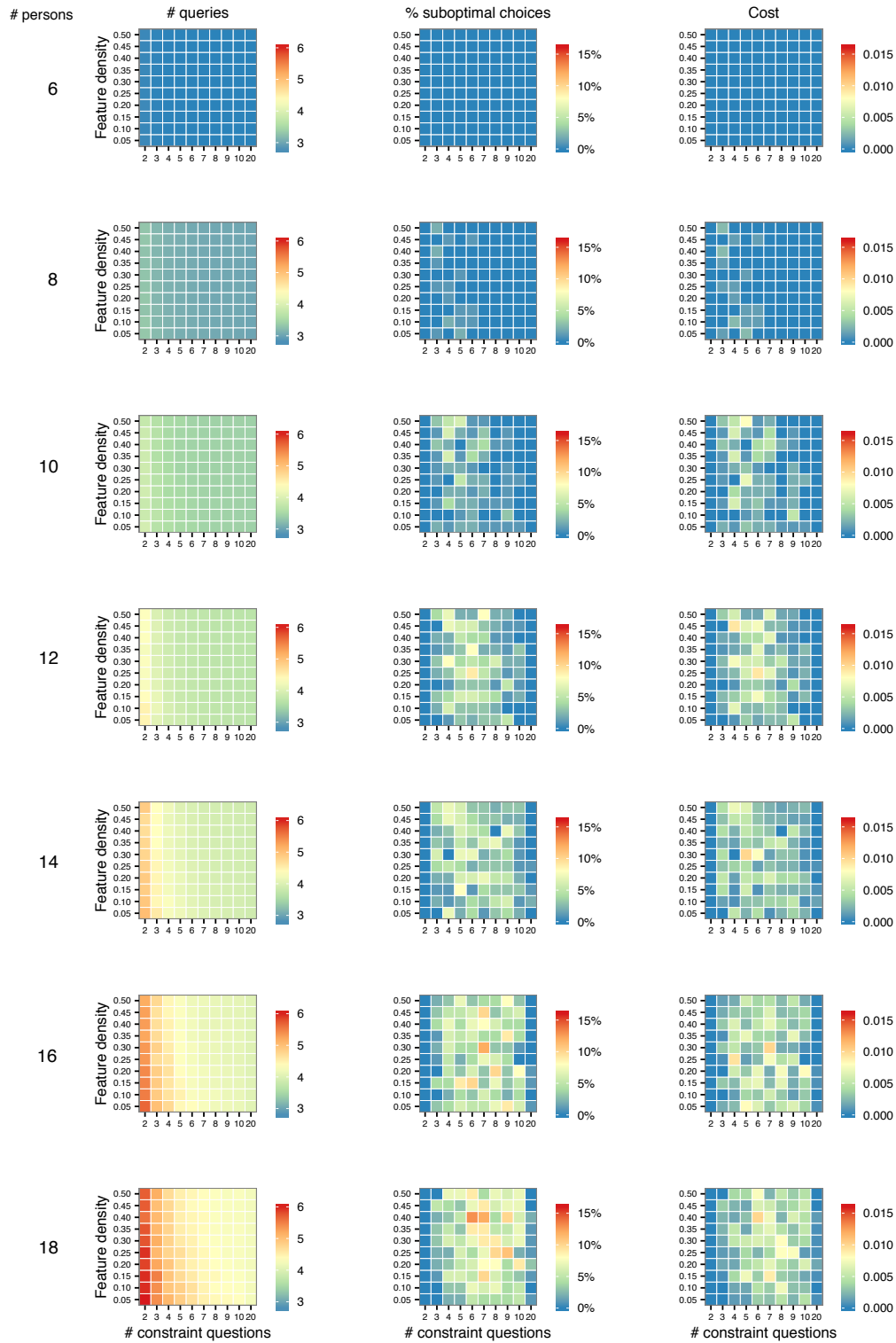


Figure C3. Detailed simulation results for many Person Game environment simulations, with correlated features. “Correlated features” means that the first constraint feature is generated with the average feature density specified, and each subsequent constraint feature has a .5 probability of also being generated in this way (i.e., of being generated independently of all previous features) and a .5 probability of being generated by randomly changing the feature vector of the previous feature, flipping each bit in that vector with probability .2. Other details of the plot are the same as explained in Figure C1.

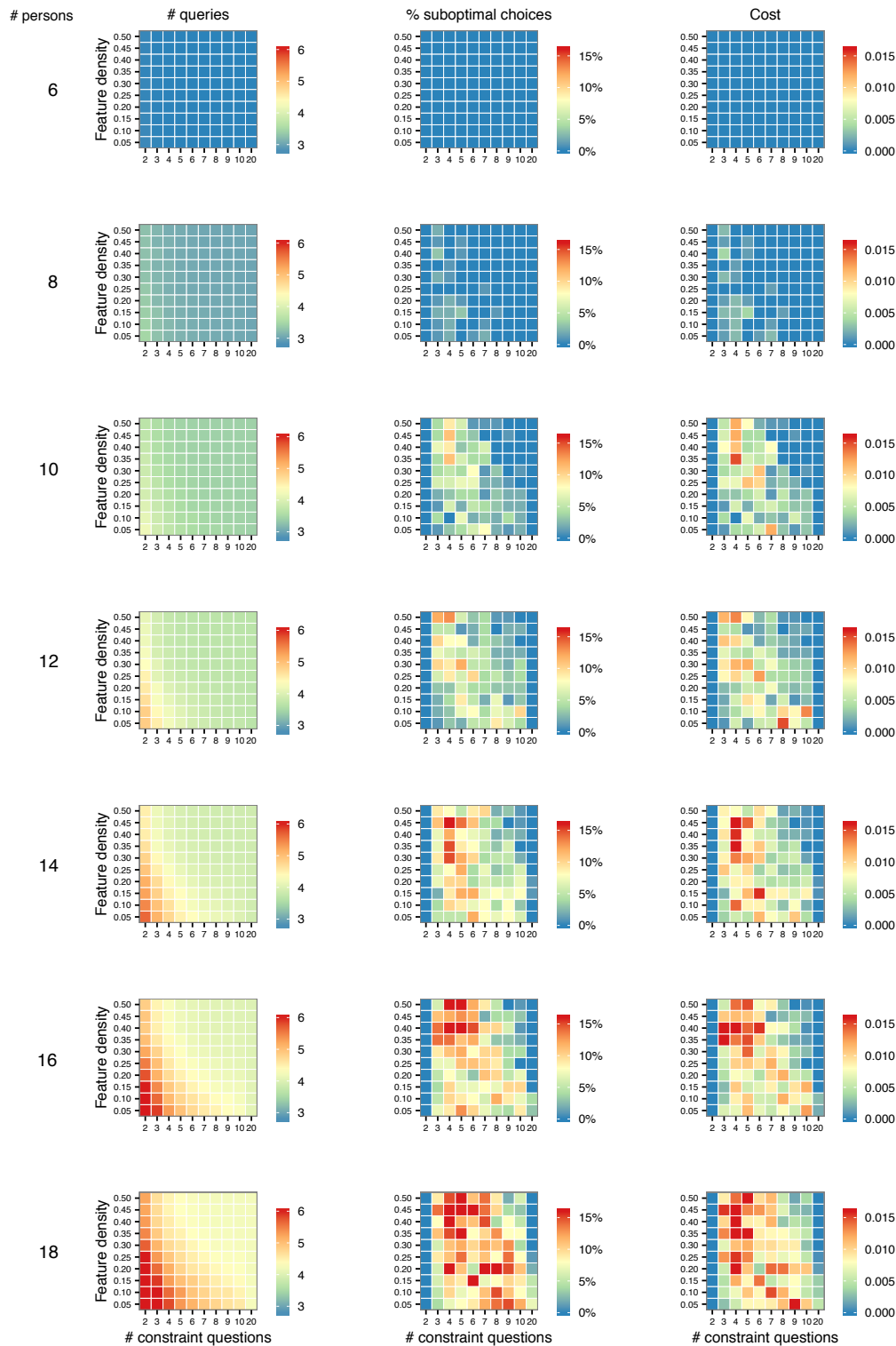
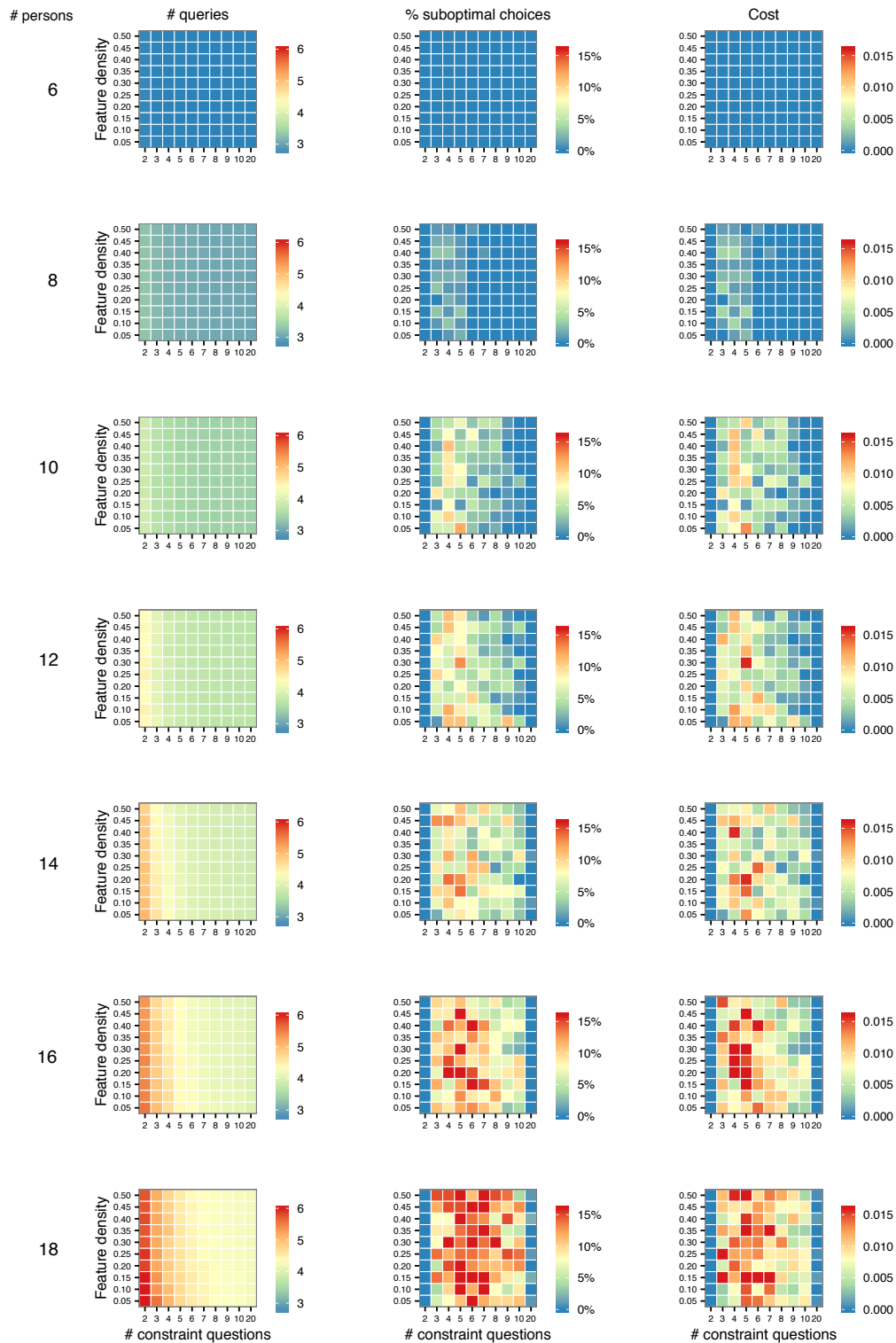


Figure C4. Detailed simulation results for many Person Game environment simulations, with correlated diverse features. “Correlated diverse features” denotes a combination of the diverse feature and correlated feature methods. Other details of the plot are the same as explained in Figure C1. This feature generation method, which we believe is the most naturalistically plausible among the four methods, led to the worst performance of the stepwise methods.



## Appendix A

**Theorem 1:** *If likelihoods are deterministic, then a query's entropy is the same as its expected information gain; i.e., if  $\text{ent}(Q|H) = 0$ , then  $eu_{IG}(Q) = \text{ent}(Q)$ .*

*Proof.* Let  $I(H; Q)$  denote the mutual information (Cover & Thomas, 1991) between the unknown hypothesis and the question (query) which has not yet been answered,

$$I(H; Q) = \text{ent}(H) - \text{ent}(H|Q), \quad (\text{B1})$$

where  $\text{ent}(H)$  denotes the prior entropy of the hypotheses, and  $\text{ent}(H|Q) = \sum_{j=1}^m P(q_j) \text{ent}(H|q_j)$  is the conditional entropy of  $H$  given  $Q$ . Thus  $I(H; Q)$  is the expected information gain of the question  $Q$ . By symmetry of mutual information (Cover & Thomas, 1991),

$$I(H; Q) = I(Q; H) = \text{ent}(Q) - \text{ent}(Q|H) \quad (\text{B2})$$

where  $\text{ent}(Q)$  denotes the prior entropy of the question and  $\text{ent}(Q|H)$  denotes the conditional entropy of  $Q$  given  $H$ ,  $\text{ent}(Q|H) = \sum_{i=1}^n P(h_i) \text{ent}(Q|h_i)$ . Now, in any domain with deterministic likelihoods, if the hypothesis  $h_i$  is known, then the answer to every question (the result of every query) is known with certainty. Therefore, in any search task wherein the likelihoods are deterministic, it holds that  $\text{ent}(Q|h_i) = 0$  for all  $h_i$  and all questions  $Q$ , and therefore  $\text{ent}(Q|H) = \sum_{i=1}^n P(h_i) \cdot 0 = 0$ . Thus, if feature likelihoods are deterministic,

$$eu_{IG}(Q) = I(Q; H) = \text{ent}(Q) - \text{ent}(Q|H) = \text{ent}(Q), \quad (\text{B3})$$

and the maximum entropy question is also the maximum information gain question.

**Theorem 2:** *In the Person Game, a question's expected impact,  $eu_{\text{Imp}}(Q)$ , is equal to  $4P(q_{\text{yes}})P(q_{\text{no}}) = 4P(q_{\text{yes}})[1-P(q_{\text{yes}})]$ . Consequently, the split-half strategy invariably selects the highest impact question in the Person Game.*

*Proof.* Impact quantifies the usefulness of a datum as the absolute change in belief from the prior to the posterior probability, summed over all hypotheses (Equation 2). In the Person Game with  $n$  persons, the prior probability before asking a question is  $1/n$  for all hypotheses (persons in the current set), the posterior probability for all remaining items after a positive answer is  $1/n_{\text{yes}}$ , and the posterior probability for all items eliminated by the question is 0. Accordingly, for each person remaining in the set the change in belief is  $1/n_{\text{yes}} - 1/n$  and for each person eliminated the change is  $1/n$ . The impact of a positive answer is then given by

$$\begin{aligned} u_{\text{Imp}}(q_{\text{yes}}) &= \sum_{i=1}^n |P(h_i|q_{\text{yes}}) - P(h_i)| & (B4) \\ &= n_{\text{yes}} \left( \frac{1}{n_{\text{yes}}} - \frac{1}{n} \right) + n_{\text{no}} \frac{1}{n} \\ &= 1 - P(q_{\text{yes}}) + P(q_{\text{no}}) \\ &= 2P(q_{\text{no}}) \end{aligned}$$

and similarly the impact of a negative answer is given by

$$\begin{aligned} u_{\text{Imp}}(q_{\text{no}}) &= \sum_{i=1}^n |P(h_i|q_{\text{no}}) - P(h_i)| & (B5) \\ &= n_{\text{no}} \left( \frac{1}{n_{\text{no}}} - \frac{1}{n} \right) + n_{\text{yes}} \frac{1}{n} \\ &= 1 - P(q_{\text{no}}) + P(q_{\text{yes}}) \\ &= 2P(q_{\text{yes}}). \end{aligned}$$

The expected impact of a question,  $eu_{\text{Imp}}(Q)$ , in the Person Game is therefore

$$\begin{aligned} eu_{\text{Imp}}(Q) &= P(q_{\text{yes}})u_{\text{Imp}}(q_{\text{yes}}) + P(q_{\text{no}})u_{\text{Imp}}(q_{\text{no}}) & (B6) \\ &= 4P(q_{\text{yes}})P(q_{\text{no}}) \end{aligned}$$

$$= 4P(q_{\text{yes}}) - 4P(q_{\text{yes}})^2 .$$

Note that the expected impact of a question in the person game does not depend on the number of items  $n$ , but only on  $P(q_{\text{yes}})$ , where  $P(q_{\text{yes}}) = n_{\text{yes}}/n$  and  $P(q_{\text{no}}) = 1 - P(q_{\text{yes}})$ . Importantly, this function is a two-degree polynomial and therefore continuous and continuously differentiable; it has its maximum at  $P(q_{\text{yes}}) = .5$  and decreases monotonically from that point in both directions (Figure 2).

**Theorem 3:** *In the Person Game, the split-half heuristic maximizes stepwise expected information gain.*

*Proof.* A binary question has maximum entropy if  $P(q_{\text{yes}}) = P(q_{\text{no}}) = .5$ ; the closer  $P(q_{\text{yes}})$  is to .5, the higher the entropy in that binary question. Moreover, the entropy of a binary question decreases monotonically and symmetrically as  $P(q_{\text{yes}})$  moves away from .5. Therefore, the most splithalfy question is also the maximum entropy question. In the Person Game, likelihoods are deterministic, and all items are equally probable. (The fact that likelihoods are deterministic means that once the target is known, there is no uncertainty about any of the features. For instance, each person either has glasses or does not have glasses. In other words,  $\text{ent}(Q|H) = 0$ .) Therefore Theorem 1 applies, and the maximum entropy question is also the highest information gain question. Thus the split-half heuristic identifies the maximum information gain question.

**Theorem 4:** *In the Person Game, where there are  $n$  items (hypotheses) remaining, all informative questions have expected probability gain  $1/n$ .*

*Proof.* The  $n$  items are equally probable a priori, so the probability of guessing correctly before asking a question is  $1/n$ . The probability of guessing correctly after obtaining a positive answer is  $1/n_{\text{yes}}$ , and the probability of guessing correctly after obtaining a negative answer is  $1/n_{\text{no}}$ . Accordingly, for any question  $Q$  the usefulness of positive and negative answers, respectively, is given by:

$$u_{\text{PG}}(q_{\text{yes}}) = \max_{1 \leq i \leq n} P(h_i | q_{\text{yes}}) - \max_{1 \leq i \leq n} P(h_i) = \frac{1}{n_{\text{yes}}} - \frac{1}{n} \quad (\text{B7})$$

$$u_{\text{PG}}(q_{\text{no}}) = \max_{1 \leq i \leq n} P(h_i | q_{\text{no}}) - \max_{1 \leq i \leq n} P(h_i) = \frac{1}{n_{\text{no}}} - \frac{1}{n} \quad (\text{B8})$$

For computing the expected probability gain of a query  $Q$ , the usefulness of each possible outcome is weighted by its probability of occurrence, yielding<sup>2</sup>

$$\begin{aligned} eu_{\text{PG}}(Q) &= P(q_{\text{yes}})u_{\text{PG}}(q_{\text{yes}}) + P(q_{\text{no}})u_{\text{PG}}(q_{\text{no}}) \quad (\text{B9}) \\ &= \frac{n_{\text{yes}}}{n} \left( \frac{1}{n_{\text{yes}}} - \frac{1}{n} \right) + \frac{n_{\text{no}}}{n} \left( \frac{1}{n_{\text{no}}} - \frac{1}{n} \right) \\ &= \frac{1}{n} \end{aligned}$$

The result of this analysis is that *every* informative question has the same probability gain, specifically  $1/n$ , *irrespective* of the proportion of objects with the feature.

**Theorem 5:** *If there are fewer than eight items in a Person Game environment, then the stepwise methods (i.e., the split-half heuristic, maximum expected information gain, and maximum expected impact) invariably select the first question with minimal expected path length, up to a possible tie.*

*Proof.* To prove the theorem, it suffices to show, for any number of items  $n \leq 7$  and for any two questions  $Q_1$  and  $Q_2$  where  $Q_1$  is more splithalfy than  $Q_2$ , that the least efficient tree starting with  $Q_1$  is as efficient (i.e., has as short an expected path length) as the most efficient tree starting with  $Q_2$ . To be thorough, we do this systematically, considering person games of size 3 through 7 in order.

If there are three items, every question is a name question, and there is just one possible tree (up to isomorphism), so the split-half heuristic and all other strategies are optimal and equivalent.

<sup>2</sup> We are indebted to Gudny Gudmundsdottir for this result.



If there are four items, there are two possible trees: a tree with a 1:3 initial question, and a tree with a 2:2 initial question. The tree with the 1:3 initial split requires 2.25 questions on average, whereas the tree with a 2:2 initial split requires 2 questions on average. Thus the split-half heuristic is optimal in the case of four items.

If there are five items, there are two possible initial splits, 1:4 and 2:3. Just one tree starts with the 2:3 split; it requires 2.4 questions on average. The most efficient possible tree that starts with a 1:4 split follows with a 2:2 split in case the first answer leaves 4 hypotheses; this tree requires 2.6 questions on average. Thus the split-half heuristic is optimal in the case of five items.

If there are six items, there are three possible initial splits, 1:5, 2:4, and 3:3. Three question trees begin with a 1:5 split; the most efficient among them has a 2:3 split of the group of 5 with the second question and requires 3 questions on average. Two question trees begin with a 2:4 split (according to whether the group of 4 is split 2:2 or 1:3 with the second question), and they require 2.67 and 2.83 questions on average. Only one question tree begins with a 3:3 split; it requires 2.67 questions on average. Thus it is not possible, in a Person Game with six items, for a tree that starts with a less splithalfy question to be more efficient than a tree that starts with a more splithalfy question.

If there are seven items, there are three possible initial splits, 1:6, 2:5, and 3:4. The most efficient tree that starts with a 1:6 split divides the group of 6 as 3:3 with the second question and requires 3.29 questions on average. The least efficient tree that starts with a 3:4 split has 1:2 and 1:3 splits at the second step and requires 3 questions on average. There are three trees that begin with a 2:5 split (according to whether the second question splits the group of 5 as 2:3 or 1:4, and in the latter case whether the third question splits the group of 4 as 2:2 or 1:3), and they require 3, 3.14, and 3.29 questions on average. Thus again it is not possible, in a Person Game with seven items, for a tree that starts with a less splithalfy question to be more efficient than a tree that starts with a more splithalfy question.

**Theorem 6:** *If a Person Game environment has only two constraint questions, then the stepwise methods invariably select the first question with minimal expected path length.*

*Proof.* Label the two constraint questions  $Q_1$  and  $Q_2$ . If these questions are equally splithalfy (have equal entropy) then the split-half heuristic has no preference between them, and likewise starting with either will lead to the same expected path length. Therefore assume that  $Q_1$  is more splithalfy than  $Q_2$ , and hence that the split-half heuristic starts with  $Q_1$ .

Case 1: Assume that  $Q_1$  and  $Q_2$  are logically independent, meaning the four logical cells of hypotheses they jointly define are all nonempty. Let  $T_1$  be any tree that asks  $Q_1$  first,  $Q_2$  second, and informative name questions thereafter. This tree is stepwise optimal. Now let  $T_2$  be a tree that asks  $Q_2$  first,  $Q_1$  second, and informative name questions thereafter. Because  $Q_1$  is informative regardless of the answer to  $Q_2$ , it is at least as effective as any name question, and therefore  $T_2$  has minimal expected path length among all trees starting with  $Q_2$ . Now observe that  $T_1$  and  $T_2$  must have the same expected path length, because for any value of the target, these two trees put the player in the same state after two questions. In other words, asking  $Q_1$  first and  $Q_2$  second yields the same result as the other way around. Similar arguments show that  $T_1$  has expected path length less than or equal to any tree starting with a name question. Therefore  $T_1$  is globally optimal.

Case 2: If instead  $Q_1$  and  $Q_2$  are not logically independent, then one of their four logical cells is empty. (If two or more cells are empty, then one question is uninformative or the questions are equivalent.) Without loss of generality, assume there are no targets with both questions negative. Let  $T_1$  be any tree that asks  $Q_1$  first,  $Q_2$  second if  $Q_1$  is positive, and informative name questions in all other situations. This tree is stepwise optimal. Let  $T_2$  be a tree that asks  $Q_2$  first,  $Q_1$  second if  $Q_2$  is positive, and informative name questions in all other situations. The rest of the argument proceeds as in Case 1:  $T_2$  has minimal expected path length among all trees starting with  $Q_2$ , and  $T_1$  has the same expected path length. Therefore  $T_1$  is globally optimal.

In both of the above cases, we have proven that there exists a stepwise-optimal tree that is globally optimal. In fact, one can use the same approach to prove that all stepwise-optimal trees are globally optimal. The full proof involves cases where the stepwise approach considers asking a name question on the second step to be as good as asking  $Q_2$ , because one or more of the four logical cells mentioned above has only one member. We omit these details for sake of space and because they follow similar arguments to the ones above.

Finally, note that this proof (or a simple extension thereof) yields a stronger result than stated in the theorem: If there are only two constraint questions, then expected path length can be minimized by asking either of them as the first question.

**Theorem 7:** *In the Person Game, the expected probability gain of any question strategy (tree  $T$ ) is given by  $eu_{PG}(T) = (r - 1)/n$ , where the (nonempty) leaf nodes of the tree are indexed by  $j=1, 2, \dots, r$ , and there are  $n$  persons (items).*

*Proof.* From Theorem 8 (see below) we have that the probability gain of any Generalized Person Game tree is  $eu_{PG}(T) = \sum_{j=1}^r P(h_j^*) - \max_{1 \leq j \leq r} P(h_j^*)$ . In the (standard) Person Game, all properties of the Generalized Person Game also hold, but there is the additional constraint that the prior probability of each hypothesis is  $1/n$ . Thus  $eu_{PG}(T) = \sum_{j=1}^r \frac{1}{n} - \frac{1}{n} = \frac{r-1}{n}$ .

**Theorem 8:** *In the Generalized Person Game, the probability gain of any question strategy (tree  $T$ ) is given by  $eu_{PG}(T) = \sum_{j=1}^r P(h_j^*) - \max_{1 \leq j \leq r} P(h_j^*)$ , where the terminal nodes of the tree are indexed by  $j=1, 2, \dots, r$ , and  $h_j^*$  is the hypothesis in node  $j$  that has highest prior probability.*

*Proof.* First, consider the probability gain of asking a single question  $Q$ , which is equivalent to a tree of depth  $k=1$ . Let  $h_Y^*$  denote the hypothesis with the highest prior probability among those that will yield a positive answer to the question, and let  $h_N^*$  denote the hypothesis with the highest prior probability among those that will yield a negative

answer. In each case, if two or more hypotheses tie for the highest probability, then the player can choose arbitrarily among those hypotheses. Therefore the searcher will guess correctly if the target is either  $h_Y^*$  or  $h_N^*$ , implying the ex ante expected accuracy equals the sum of their prior probabilities,  $P(h_Y^*) + P(h_N^*)$ . Furthermore, the expected accuracy without asking the question equals the probability of whichever hypothesis has the largest prior probability, which is necessarily  $h_Y^*$  or  $h_N^*$ . Therefore the expected probability gain of the question is

$$\begin{aligned} eu_{PG}(Q) &= P(h_Y^*) + P(h_N^*) - \max\{P(h_Y^*), P(h_N^*)\} \\ &= \min\{P(h_Y^*), P(h_N^*)\}. \end{aligned} \quad (B10)$$

Now consider an arbitrary tree  $T$ , with depth  $k$  and with  $r$  leaf nodes labeled  $t_1, t_2, \dots, t_r$ . For the purpose of guessing the target hypothesis, the process of following this question tree is equivalent to asking a single  $r$ -valued question, with the leaf node that results from the tree corresponding to the answer to that question. Thus,  $t_1, t_2, \dots, t_r$  can be thought of as answers to the question  $T$ , paralleling the terminology elsewhere in this paper of a single question  $Q$  with answers  $q_{yes}$  and  $q_{no}$ . The analysis above for the probability gain of a single question can easily be extended, by defining  $h_j^*$  to be a hypothesis with maximal prior probability among those that will lead to answer (i.e., leaf node)  $t_j$ . This hypothesis will be the searcher's guess if  $t_j$  is the observed outcome, and therefore the guess will be correct iff the target is  $h_j^*$  for some  $j$ . Thus the expected accuracy from using the tree is  $\sum_{j=1}^r P(h_j^*)$ , and the expected probability gain of the tree is

$$eu_{PG}(T) = \sum_{j=1}^r P(h_j^*) - \max_{1 \leq j \leq r} P(h_j^*). \quad (B11)$$

To demonstrate this result more rigorously, let  $L_j$  denote the set of hypotheses at leaf node  $t_j$ , i.e., all hypotheses that are not eliminated by the sequence of question(s) and answer(s) that lead to that node. For instance, if  $h_3$  and  $h_5$  remain as the only possible

hypotheses at leaf node  $t_1$ , then  $L_1 = \{h_3, h_5\}$ . The classification accuracy that will be obtained at leaf node  $t_j$  is the maximum posterior probability at that point,  $\max_{1 \leq i \leq n} P(h_i | t_j)$ . Thus, the expected classification accuracy from the tree is

$$P(\text{correct}|T) = \sum_{j=1}^r P(t_j) \max_{1 \leq i \leq n} P(h_i | t_j). \quad (\text{B12})$$

The posterior probability  $P(h_i | t_j)$  can be derived via Bayes' rule, i.e.  $P(h_i | t_j) = P(t_j | h_i) P(h_i) / P(t_j)$ . In the Person Game, likelihoods are deterministic, so  $P(t_j | h_i) = 1$  for all  $h_i \in L_j$  and  $P(t_j | h_i) = 0$  for all  $h_i \notin L_j$ . Accordingly, Bayes' rule simplifies and Equation 19 can be rewritten as

$$\begin{aligned} P(\text{correct}|T) &= \sum_{j=1}^r P(t_j) \max_{h_i \in L_j} \frac{P(h_i)}{P(t_j)} \\ &= \sum_{j=1}^r \max_{h_i \in L_j} P(h_i). \end{aligned} \quad (\text{B13})$$

In other words, to calculate the classification accuracy of a tree, one needs only to sum the highest-prior-probability hypothesis (person) at each possible leaf node. The probability gain of the tree is its classification accuracy minus the classification accuracy that could be achieved based on the prior probabilities alone:

$$\begin{aligned} eu_{\text{PG}}(T) &= \sum_{j=1}^r P(h_j^*) - \max_{1 \leq i \leq n} P(h_i) \\ &= \sum_{j=1}^r P(h_j^*) - \max_{1 \leq j \leq r} P(h_j^*). \end{aligned} \quad (\text{B14})$$

## Appendix B

**The Number of Search Trees**

Let  $m$  denote the total number of (binary) queries ( $Q_1$  through  $Q_m$ ) in a binary sequential-search task, and let  $k$  index the time step starting at zero for the first query (so that  $0 \leq k < m$ ). We derive here the number  $N_m$  of full question trees for such a task.<sup>3,4</sup> Note first that the number of branches at each step in a tree (i.e., decision points at which a question must be selected for the next step) grows exponentially with the number of questions already asked. After the first question there are 2 branches (one for each possible outcome of that question), after the second question there are 4 branches (one for each of the 4 possible combinations of outcomes of the first and second questions), and in general there are  $2^k$  branches at step  $k$  (after  $k$  questions have been asked). At each branch, at each time step  $k$ , there are  $m-k$  questions available. Because questions can be assigned independently to the different branches, the total number of joint assignments of questions to branches at step  $k$ , conditioned on the questions assigned at all previous steps, is equal to  $(m-k)^{2^k}$ .

For instance, if  $m = 3$ , there are 3 possible root nodes (initial questions) at step 0. In each case, the first question may result in one of two outcomes (2 branches). On step 1, there are 2 remaining queries to choose from on each branch, so there are  $2^2 = 4$  possible question assignments. After this, there is only one question left on each branch. So, for  $m = 3$  queries, there are  $N_3 = 3 \times 4 = 12$  full question trees. If  $m = 4$ , there are 4 possible root nodes (initial questions) at step 0. At step 1 there are 3 queries left to choose from on each of the two branches, so there are  $3^2 = 9$  possible question assignments. On step 2, there are 4 branches,

---

<sup>3</sup> By full question tree, we mean a tree that exhausts all available questions even if the target hypothesis has been identified in fewer than  $n$  queries. The number of practically distinct question trees (i.e., ones that produce different outcomes for the set of hypotheses under consideration) will be less than the number of full question trees in many realistic tasks.

<sup>4</sup> Binary question trees should not be confused with the concept of binary search trees in computer science, a class of data structures (e.g., Knuth, 1971).

each with 2 possible queries to choose from, and thus  $2^4 = 16$  possible question assignments. After this, there is no degree of freedom, since only one question is left at each branch. So, in total there are  $N_4 = 4 \times 9 \times 16 = 576$  full question trees.

In general, the number of full question trees in a task with  $m$  available questions is equal to the product over time steps of the number of joint question assignments at each step (conditioned on the assignments at the previous steps):

$$N_m = \prod_{k=0}^{m-1} (m-k)^{2^k}. \quad (\text{A1})$$

For  $m$  equal to 1 through 6, this formula yields 1, 2, 12, 576, 1658880, and 16511297126400 unique question trees. Algebraic rearrangement allows the number of trees to be re-expressed as

$$\begin{aligned} N_m &= 2^{\sum_{k=0}^{m-1} 2^k \log_2(m-k)} \\ &= 2^{\sum_{j=1}^m 2^{m-j} \log_2 j} \\ &= 2^{2^m \sum_{j=1}^m \frac{\log_2 j}{2^j}} \\ &= 2^{(m+c_m)} \end{aligned} \quad (\text{A2})$$

where

$$c_m = \log_2 \sum_{j=1}^m \frac{\log_2 j}{2^j}. \quad (\text{A3})$$

The series in Equation A3 converges, and numerically one finds that  $\lim_{m \rightarrow \infty} (c_m) \approx -.449$ . Therefore the number of trees behaves asymptotically as

$$N_m \approx 2^{2^{(m-.449)}} \quad (\text{A4})$$

which grows faster than any exponential function of  $m$ .

### Finding the Most Efficient Tree

Given a set of  $m$  possible binary queries that are jointly sufficient to uniquely identify any of the hypotheses, the most efficient binary question tree is defined as the tree with the

shortest expected path length (alternate definitions are possible, of course). Because the number of possible trees grows super-exponentially in the number of possible queries, exhaustive search is not feasible except for very small  $m$ . For moderate values of  $m$ , dynamic programming can be used to identify the optimal tree. Specifically, one can start at step  $m - 2$  and determine the optimal question (from the two remaining questions) for all  $m(m-1)2^{m-3}$  possible states of the game at that step, which are defined by choices for the first  $m - 2$  queries (ignoring order) together with the possible answers to those questions. (Note that many of these combinations are impossible, because they do not correspond to any hypothesis, and thus the actual memory requirements of this algorithm are in general considerably smaller.) Then one can do the same for step  $m - 3$ , determining the optimal subtree for all possible game states at that step, based on the results for step  $m - 2$ . In general, for each game state at step  $k$ , one can use the conditional probabilities for the answers to all unasked questions, together with the optimal subtrees rooted at step  $k + 1$  and their expected path lengths, to determine the optimal subtree and expected path length for the state in question. Proceeding by backward induction yields the optimal tree for the whole task (i.e., rooted at step 0). For instance, applying this method to our example Person Game environment (Table 1) yields the globally optimal tree (Figure 3b, top row).



### Appendix C

This appendix describes the simulations of the performance of the split-half heuristic in greater detail than the text. (Recall that a searcher who selects questions so as to maximize stepwise expected impact (Theorem 2) or information gain (Theorem 3) always asks the same questions as the split-half heuristic.) The simulations were designed to check whether our example scenario (Table 1) is unusual, or whether the split-half strategy is suboptimal in a wide range of Person Game environments. These simulations explore various parameters of interest, such as how features are generated, and the number of features (e.g., of possible questions) in the environment.

In the simulations, random Person Game environments were constructed as a function of four parameters: type of features (explained below), number of persons (hypotheses), number of constraint questions, and average feature density. For each randomly generated Person Game environment, we checked the following: the expected number of questions required if the optimal strategy, including the best first question, is used; whether the split-half heuristic selects the best first question; and the cost (if any, in terms of average total number of questions required) of using the split-half heuristic. If more than one question ties for being most splithalfy, we assume that the split-half heuristic will select among those questions with equal probability. In case the most splithalfy first question(s) do not correspond to the optimal first questions, we calculate the cost of using the split-half heuristic as the average additional number of questions required, if the first question is selected at random among the maximally splithalfy first questions, as opposed to the optimal first question. For the second and subsequent questions, the simulation code presumes that the optimal question will be asked given what has been learned so far. (This means that simulation results could understate the cost of using the split-half heuristic, especially for environments with large numbers of items.)

We illustrate the environment-generating procedure with an example with 10 persons, 3 constraint questions, and feature density .4. This specifies construction of a random Person Game environment with 10 items, a name feature for each item (a feature possessed by only that item), and 3 additional constraint features. For the baseline (“regular features”) type of features, the procedure to construct a constraint feature is as follows: for each of the 10 items, with IID (independent and identically distributed) probability .4, the person has that particular feature. Keep the resulting feature if it is informative (i.e., it is possessed by at least one of the persons, and not by all of the persons), and is not a name feature (or the complement of a name feature). Otherwise repeat this generation process until a suitable feature is found. Repeat this process until all three constraint features (questions) have been generated.

Whereas most of the above parameters are self-explanatory, the type of feature parameter should be explained on its own. Each feature is described by a binary vector with length  $n$ , where there are  $n$  items (faces) in the game. In each case, a name feature was created for each of the items. Additional constraint features were created, as described below. The types of constraint feature generation methods were:

- *regular features*: set each item in the feature vector to present (value 1) with the specified target feature density probability, independent of each other item in the feature vector. Repeat this process, independently, for all constraint features.
- *diverse features*: for each feature, obtain a target feature density by taking a sample from a Beta distribution with mean target feature density, and parameters  $\alpha=s*\text{targetFeatureDensity}$ ,  $\beta=s*(1-\text{targetFeatureDensity})$ , where  $s=2$ .<sup>5</sup> Generate a candidate constraint feature by setting each item in the feature vector to present (value

---

<sup>5</sup> When setting the  $\alpha$  and  $\beta$  parameters, use of  $s=2$  in the equation  $\alpha=s*\text{targetFeatureDensity}$ ,  $\beta=s*(1-\text{targetFeatureDensity})$  is somewhat arbitrary. Our choice of  $s=2$  entails that if  $\text{targetFeatureDensity}=.5$ , each feature’s target density will be selected from a uniform distribution. Larger  $s$  would reduce the variability in the density of features generated. (In the limiting case, where  $s$  is unboundedly large, the diverse feature generation method reduces to the regular features generation method.) Smaller  $s$  will result in a tendency to select features with extreme (e.g. 1 or 0) density values.

- 1) with that sampled probability. Repeat this process (with different samples from the beta distribution) independently for each constraint feature.
- *correlated features*: obtain the first constraint feature as specified in the “regular features” feature construction method. For each subsequent constraint feature, flip a fair coin; if the coin comes up heads, use the same method as for the initial feature. If the coin comes up tails, take the most recently generated constraint feature, and flip each value in it with independent probability .2.
  - *correlated diverse features*: for the first constraint feature, use the “diverse features” construction method. For each subsequent constraint feature, flip a fair coin. If the coin comes up heads, repeat the “diverse features” construction method. If the coin comes up tails, take the most recently generated constraint feature, and flip each value in it with independent probability .2.

Note that the feature density parameter refers to the generating probability, and will differ to some extent from the resulting sampled mean feature density, especially in the case of very small size Person Game environments and low feature densities, due to the requirement that each constraint question be an informative non-name question. There are some differences according to feature generation method, as well; for instance, with the correlated features method, the flipping procedure has the effect of moving the expected resulting feature density towards .5. Although generation of duplicate constraint features (questions) is infrequent, the code did not require that each constraint question be unique.

Each combination of the following parameters was simulated: 4 types of features {regular features, diverse features, correlated features, diverse correlated features}, number of persons {6, 8, 10, 12, 14, 16, 18}, number of constraint questions {2, 3, 4, 5, 6, 7, 8, 9, 10, 20}, and target feature density {.05, .10, ..., .50}. Each combination of parameter values was used to create 100 random Person Game environments.

Figure 5 reports the average proportion of instances in which the split-half strategy selects a suboptimal first question, as a function of the type of feature generation method and the number of items (faces) in the random Person Game environments. This figure averages across all of the different numbers of constraint questions and all different feature density values. For every feature construction method, the probability of the split-half heuristic choosing a suboptimal first question increases with the number of faces in the game. Correlated features and heterogenous density both make split-half more likely to be suboptimal.

Figure C1 explains the axes of each of the main figures. Figures C1 through C4 show detailed results, with the performance of the split-half strategy by feature construction method, number of items (faces), number of constraint questions, and feature density. Figure C1 is for the regular feature method; Figure C2, diverse features; Figure C3, correlated features; Figure C4, diverse correlated features. Figures 5a and 5b (in the main text) show aggregate results for the frequency that the split-half strategy was suboptimal, according to the number of items (Figure 5a) and the number of constraint questions (Figure 5b), aggregating across other simulation parameters. Performance of the split-half heuristic was typically worse to the extent that there were many items, and/or only a small number of constraint features.