A Cognitive Model of How People Make Decisions Through Interaction with Visual Displays

Xiuli Chen
School of Psychology,
University of Birmingham,UK
chenxy@bham.ac.uk*

Sandra Dorothee Starke
Electronic, Electrical and Systems Engineering,
University of Birmingham,UK
s.d.starke@bham.ac.uk

Chris Baber
Electronic, Electrical and Systems Engineering,
University of Birmingham,UK
c.baber@bham.ac.uk

Andrew Howes
School of Computer Science,
University of Birmingham,UK
HowesA@bham.ac.uk*

ABSTRACT
In this paper we report a cognitive model of how people make decisions through interaction. The model is based on the assumption that interaction for decision making is an example of a Partially Observable Markov Decision Process (POMDP) in which observations are made by limited perceptual systems that model human foveated vision and decisions are made by strategies that are adapted to the task. We illustrate the model by applying it to the task of determining whether to block a credit card given a number of variables including the location of a transaction, its amount, and the customer history. Each of these variables have a different validity and users may weight them accordingly. The model solves the POMDP by learning patterns of eye movements (strategies) adapted to different presentations of the data. We compare the model behavior to human performance on the credit card transaction task.

ACM Classification Keywords
H.1.2 User/Machine Systems: Human factors, Human information processing

Author Keywords
cognitive modeling; visual search; eye movements; Markov Decision Process; reinforcement learning; decision making.

INTRODUCTION
Decision makers interact with computer systems in order to gather information and evaluate options. In particular, they sometimes use information visualizations that are designed to aggregate large amounts of information into readily perceptible representations of multi-dimensional data. An important research question is how do these representations help human decision making?

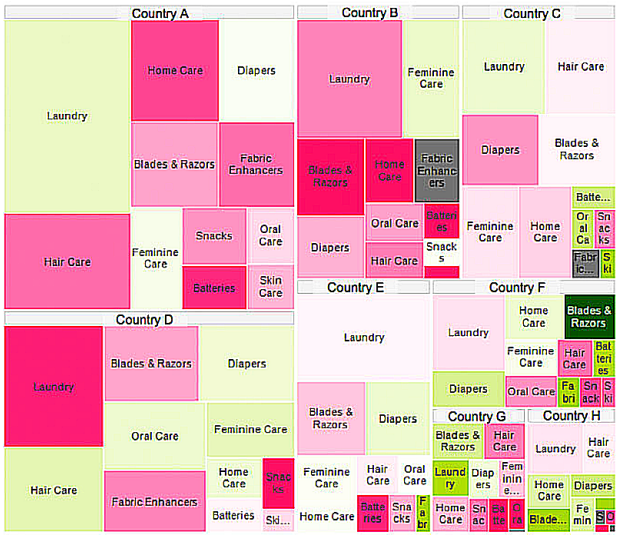

Figure 1. A color-map used by Procter & Gamble to show where their products stand in their respective markets. It shows markets in which products compete and their relative share (red indicating low market share and green indicating high market share). The color-map is used to help managers understand what is going on in the business, and to decide what to do about it. Figure from Harvard Business Review (https://hbr.org/2013/04/how-p-and-g-presents-data).

An example of a typical decision problem is deciding where and whether to make business investments. One of the visualizations that might be employed to support such a task is the color-map visualization illustrated in Figure 1. This visualization represents the relative share of Proctor and Gamble products in a range of markets in which they compete; market size is represented by the size of the rectangle and market share is represented by color (red indicating low market share and green indicating high market share). The visualization is designed with the goal of helping managers make better and faster decisions [9]. Later in the paper we describe another

CHI 2017, May 06-11, 2017, Denver, CO, USA
ACM 978-1-4503-4655-9/17/05.
http://dx.doi.org/10.1145/3025453.3025596

decision task that involves spotting potential frauds in credit card data.

Why might one expect such visualizations to be effective? We might suppose that visualizations support the recognition of complex patterns across cues; or they may offer an intuitive encoding of meaning (e.g. green indicates a good market share); or it may be that certain colors attract attention and these can be used to highlight important features. In general, the visualization is intended to enable users to directly perceive patterns in the data and this *direct perception* is assumed to result in faster and more accurate decisions. Color-map visualizations epitomize the goal of designing visualization technologies that support higher level decision making by taking advantage of human perceptual mechanisms.

Foundational work in Human-Computer Interaction developed models of the cost structure of sensemaking [37] and focused on how the design of visualizations impacts the cost of data extraction. There are likely to be differences in extraction costs (e.g. decision time) for different visualizations. In related work in intelligence analysis [1], it has been shown how users engage in parallel, overlapping explorations of data and often work with minimal and sketchy frames to explain these data. Together these contributions suggest that people will make use of a subset of the available information, partly in an effort to reduce the cost associated with information access and partly as a result of the incremental construction of a model of the data.

In the current paper, we propose an integrative cognitive model of how the process of interaction with visualizations supports decision making. It is an attempt to explain how display design can contribute to reducing information access cost and therefore efficient decision making. A key feature of the model is that the predicted user strategies are an emergent consequence of the cost structure of the task, rather than a model assumption. The model builds on a number of recent threads in Human-Computer Interaction and decision making research. These threads are described below.

1. *Partially Observable Markov Decision Processes* (POMDP). A POMDP is a mathematical framework for modeling sequential decision problems in which (1) only incomplete observations of the state of the world can be made, and (2) the transitions between states are stochastic. These properties are important for modeling interactive decision making because of the partial and stochastic nature of human information gathering. For example, using the visualization in Figure 1 to acquire information about large markets, or markets with low share, requires multiple fixations to bring high resolution foveated vision to bear on areas of interest. Each fixation retrieves only a little (partial) information and it does so with some frequency of error (see *Active Vision* below) and is therefore stochastic.

2. *Active vision*. While foveated vision is very accurate, peripheral vision is much less so [23]. As a consequence there is uncertainty as to which, and whether, objects in peripheral vision are detected. This uncertainty limits the extent to which visualizations in HCI can support direct perception. As a consequence multiple eye movements and fixations are often required to build a belief about what is on the display that is sufficient to guide action. Eye movements are, therefore, actively recruited in order to solve the user's decision task, not to build a complete model of the display. Recent work shows how color, shape, and size have different consequences for visual search of displays [23]. We used the constraints proposed in [23] to define the POMDP's observation function.

3. *Machine learning*. Human eye movements can be predicted by solving a POMDP, bounded by a model of human vision, with machine learning [6, 7]. Machine learning is used to find a strategy for choosing actions given each belief about the state of the world, where beliefs about the world are built through repeated fixations. For any POMDP there are one or more optimal strategies that choose the actions that maximize the average utility of interaction. In interactive decision tasks, the problem is to find a strategy for moving the eyes so as to generate good estimates of the task-relevant displayed state and, therefore, best possible decisions. We might expect that good visualizations facilitate efficient strategies by increasing the rate or reliability with which the human eye gathers information.

4. *Decision strategies*. Behavioral evidence suggests that people use simple but general rules (e.g., search rules, stopping rules and decision rules) in order to make good decisions quickly [16, 15, 14]. Other evidence suggests that people acquire decision strategies that are highly adapted to the cost structure of the task [19, 17]. The use of machine learning to derive strategies allows us to explore the extent to which *heuristics* are an emergent consequence of adaptation to the decision task specified in the POMDP and, therefore, the limits of the specified cognitive-perceptual mechanisms [20, 25].

Below, we first report the integrative model and then report a comparison of the model's behavior to human data. The data were collected using a laboratory study of human participants performing an ecologically determined task requiring a decision about whether a credit card payment might be fraudulent. In order to do this, participants were provided with information about various card transaction cues (e.g. transaction history, amount etc.) that were presented using one of two types of display. In one type of display, the information was presented numerically, and in the other, a color-map was used. We hypothesized that the color-map would enhance the human capacity to encode information using peripheral vision in accordance with the constraints investigated by [23]. Moreover, we used the model to predict the decision time, accuracy and number of fixations used for each interface. In addition, the study manipulated the cost of accessing each cue. In one condition, cues were always present on the display and in the other they were covered until individually revealed by a mouse click.

The contribution of the work is in providing a new model and empirical validation of how people make decisions through interaction with visualizations such as color-maps. The model

shows how decision making behavior is an emergent consequence of adaptation to display design. Critically, the model does not make any priori assumptions about decision strategy, rather the strategy is calculated using a machine learning algorithm. Detailed comparisons between the model's behavior and human performance metrics (including performance time and eye movements) are provided.

BACKGROUND

In this section we further review the background to the four threads (introduced above) that contribute to our work.

Partially Observable Markov Decision Process

It has been suggested by a number of authors that human visual search can be understood as the solution to a Partially Observable Markov Decision Process (POMDP) [6, 32, 7]. As we have said, the POMDP provides a mathematical framework that captures the interaction between an agent and the stochastic environment that it inhabits [22]. Specifically, the complete state of the environment is not known to the agent. Instead, the agent receives observations that are bounded by the limits of human perception and must make a decision based on partial information. Framing the model as a POMDP allows for the calculation of the optimal strategy given a theory of the human visual system.

For example, a target localization task was framed as a POMDP problem by [6]. In their task, a visual target was present in one position in a grid. To find the target, the agent used a series of fixations, one at each time step, each on one of the points in the grid. The information received at each time step was from the fixated point (with high reliability) and surrounding points (with lower reliability). The authors showed how the optimal strategy can generate human like behavior without assuming ad-hoc heuristics such as inhibition-of-return to previous locations. They also found that the optimal policy can contradict accepted ideas by sometimes preferring to look just to the side of the target, rather than directly at it, so as to gather more information with peripheral vision.

A random dots motion discrimination task was framed as a POMDP problem by [32]. In this task, primates were shown a visual display containing a group of moving dots. A fraction of the dots moved in one of two fixed directions (left or right) and other dots moved in random directions. The primates were rewarded for determining the direction in which the majority of dots moved, where the number of random dots was manipulated. Rao's (2010) model showed that primate decision times could be modeled as the solution to a POMDP that determined the threshold for switching from information gathering to decision.

Active vision

Studies of visual perception show that perceiving a pattern such as that in Figure 1 involves a complex sequence of eye movements to gather information and maximize the utility of the decision [42, 40, 18, 29]. This is a process of active vision. People move their eyes to seek out items within a visual scene that are relevant to the task or question they are engaged in. Eye movements are necessary since foveated vision covers only 1-2 degrees visual angle; the fovea has the highest density of daylight/color vision receptor cells. With increasing eccentricity, there is a sharp drop off in the density of these cells, and hence vision becomes rapidly blurred. The periphery, covering a much larger area than the fovea, still contains useful information despite the reduced acuity. It is well known that peripheral vision plays a key role in guiding eye movements during visual search [13] and that color, shape and size have different acuity functions [23].

Machine learning (for POMDP solutions)

The POMDP formalism is very general to a diverse range of problems [39, 26]. Unfortunately, the generality of POMDPs entails high computational cost to derive the control policy; research in the field of POMDP is mainly focused on approximate solutions [39]. For example, recent work has shown that Deep Q-Networks are capable of learning human-level control policies on a variety of different Atari 2600 games [26]. Simpler tasks can be solved with a combination of Bayesian inference for state estimation and Q-learning [47] for policy learning.

Decision strategies

The study of decision strategies has, in large part, been the study of decision heuristics. One prominent decision heuristic has received much attention: the take-the-best (TTB) heuristic [16]. The TTB heuristic consists of a set of rules concerning the most important aspects of information gathering and decision making: the search rule, the stopping rule and the decision rule. These requires knowing the validity of cues. Validity is the probability that the information represented by a cue will lead an observer to the correct decision. A person using TTB searches from the most useful information (with the highest validity) to the information with the lowest validity (the search rule). Crucially, information search terminates once a cue discriminates between the considered options or once all cues have been examined (the stopping rule); at which point the heuristic would choose the option favored by the discriminating cue (the decision rule).

Following [15] seminal work, showing how Take-The-Best (TTB) can describe human information gathering and decision making behaviors, a number of articles offered empirical investigations into which heuristics people choose [27, 28, 4, 3, 24]. A particular concern in this research has been whether people use Take-The-Best (TTB), which uses information highly selectively, or a heuristic such as Weighted-ADDitive (WADD) that integrates all information [36, 35]. While this debate has been conducted in the psychology literature, it is highly relevant to understanding how information visualizations are used by people, and therefore to how visualizations should be designed. If people use TTB and not WADD then they may make use of a much smaller part of the displayed information than if they use WADD. Also, it is possible that use of these heuristics integrates with the visualization design.

TASK

Before introducing our theory of decision making through interaction with visualizations, we briefly described a task, the

credit card fraud detection task, which will be used to illustrate the theory.

The task is motivated by a real-world scenario that is known to be extremely difficult for people. Credit card fraud detection analysts attempt to identify fraudulent patterns in transaction data-sets, often characterized by a large number of samples, many dimensions and online updates [8]. Despite the use of automated detection algorithms, there continue to be key roles for people to play in the analysis process. These roles range from defining and tuning the algorithms that automatic systems deploy, to triaging and screening recommendations from such systems, to contacting customers (either to query a transaction or to explain a decision). In terms of triaging and screening, we assume that an automated detection process is running and that this process has flagged a given transaction (or set of transactions) as suspicious and a user will engage in some form of investigation to decide how to respond to the flag. Based on pilot interviews and discussions with credit card fraud analysts and organizations, we believe that there are several ways in which the investigation could be performed. In some instances, the investigation could involve direct contact with the card-holder, in which the caller follows a pre-defined script and protocols that do not involve investigative capabilities. In some cases, these are passed to the analyst who needs to make a decision as to whether or not the credit card is blocked (this is the approach assumed in this paper). In this instance, the analyst would take a more forensic approach to the behavior of the card holder and the use of the card, relative to some concept of normal activity. In some cases, investigation could be at the level of transactions, in which the analyst seeks to identify patterns of criminal activity involving several cards. In this instance, the analysis would be looking for evidence of stolen details or unusual patterns of use of several cards, say multiple transactions in different locations within a short time-frame. Other functions that people can perform in the fraud detection process include: risk prioritization, fast closure of low risk cases, documentation of false positives [30], and identification of risk profiles and fraud patterns [38, 21].

In this study, we use a simplified version of fraud detection in which the task is to decide whether a transaction should be blocked (prevented from being authorized) or allowed. As shown in Figure 2, participants are provided with 9 sources of information (cues) and these are presented using one of 4 display designs (visualizations). The cues differ in the reliability with which they determine whether or not a transaction is a fraud and the participants must discover these validities with experience and decide which cues are worth using to make a decision. Further details are provided in the Experiment section.

The visualizations presented in Figure 2 do not have the full richness and complexity of the visualization used in Figure 1. In particular, our experimental interfaces did not use size to represent a dimension of the data. Our focus on color is partly to enable simple experimental designs and also because color is known to be more detectable with peripheral vision than shape [23].

THEORY AND MODEL

The theoretical claim made in this paper is that decision strategies are an emergent consequence of both the statistical properties of the environment (the experienced cue validities, different time cost for extracting information) and of the constraints imposed by human perceptual mechanisms. For a visualization task, this theory can be written precisely by formulating the decision problem as a POMDP for active vision and solving this problem with machine learning to find an optimal decision strategy (emergent heuristics).

In the resulting model, eye movement strategies and stopping rules, are an emergent consequence of the visualization and the limits of human vision [7, 43, 6]. The assumption is that people choose which cues to look at and when to stop looking at cues informed by the reward/cost that they receive for the decisions they make. Better decisions will receive higher rewards, which will reinforce good eye movement strategies.

In the following paragraphs, we report the problem formulation and model, followed by human data that tests the model predictions.

Problem formulation

We assume that the problem faced by a decision analyst can be modeled as a Partially Observable Markov Decision Process (POMDP) [22]. On each time step, the environment is in one of the possible states that is not directly observed. Instead, by interacting with the environment (taking actions), observations and reward (cost if the value is negative) are received from it. That is, the environment is partially observable. This action-observation-reward sequence happens in cycles indexed by $t = 1, 2, 3, ...$ Because the environment's state is not directly observed, actions are taken under uncertainty of the true environment state. On each time step, an action is chosen from the action space including both information gathering actions and decision actions. The action selection is dependent on the history of observations and actions.

To be more specific, the action-observation sequence is used to update the estimate of the underlying true state (call *'belief state'* below) using Bayesian inference. Q-learning [47] is then used to learn which action to do next (e.g., to gather more information or to make a decision) given the current belief state, so as to maximize the expected future reward. It does so by learning the belief-action values through simulated experience. Belief-action values are updated incrementally (learned) as reward and cost feedback is received from the interaction during the simulated experience. For example, if the model looks at cue A in a certain belief state and subsequently makes an incorrect decision, then the value of the action (look at cue A) at this belief state will decrease. With enough simulation trials, the optimal strategy will emerge and the model will take the best actions given the beliefs.

In the following sections more detail is provided about how the belief estimate and action selection work. A procedure for the computer simulation is provided in box Algorithm 1.

A formal description of our task as a POMDP is given in the following [22].

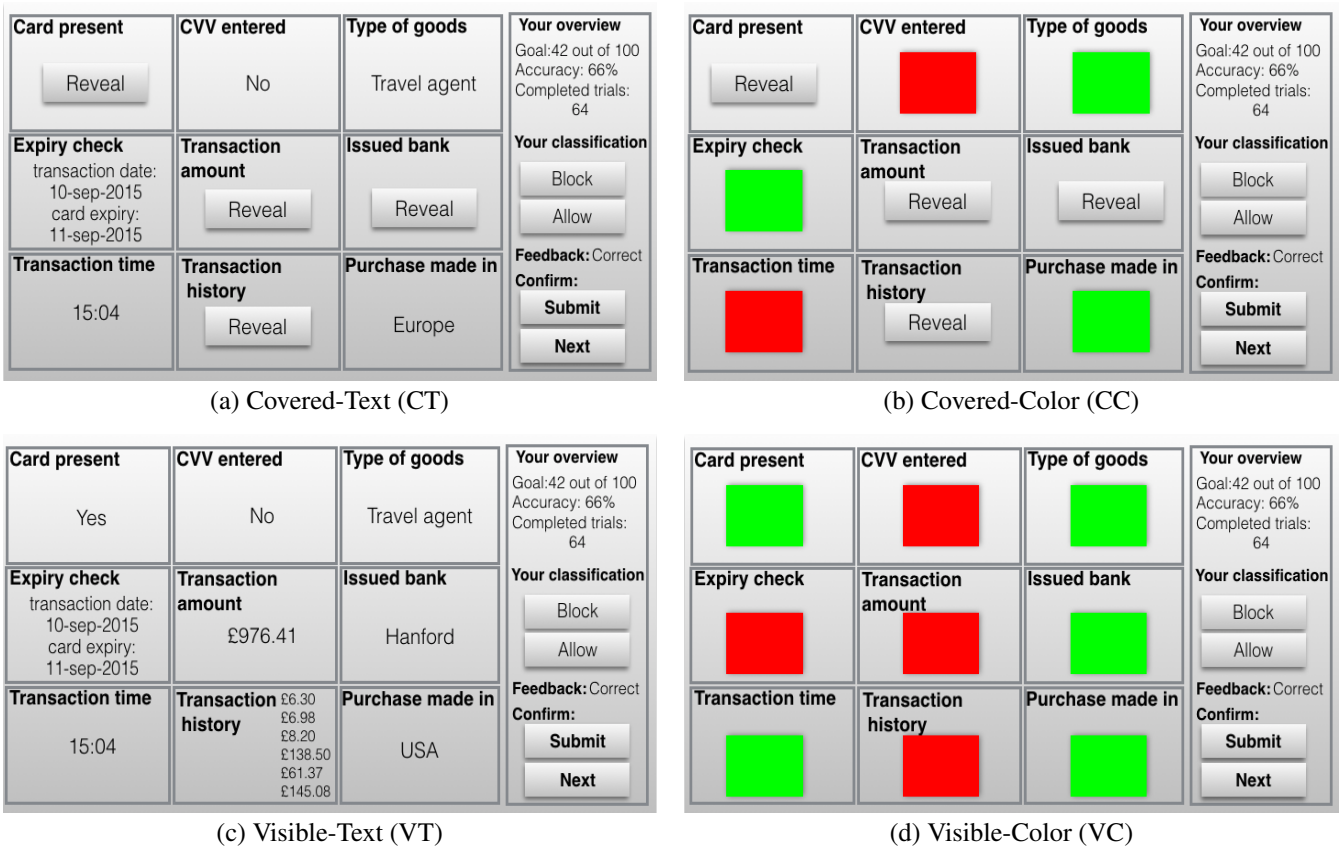

Figure 2. Four interface variants for credit card fraud detection. The information cues are represented with text (left panels) or color (right panels) and the information is either immediately available (bottom panels) or revealed by clicking the 'Reveal' buttons (top panels).

- State space S: At any time step t, the environment is in a state $s_t \in S$. A state represents a true information pattern presented on the user interface. As shown in Figure 2, nine cues associated with credit card transactions are presented on the interface. The value of each cue was discretized into two levels, representing 'fraudulent (F)' and 'normal (N)' respectively. Therefore, for example, one of the states is a 9-element vector [F,N,N,F,F,F,N,F,N], each item of which represents the value for one of the cues ('F' for fraudulent and 'N' for normal). Hence, the size of the state space is $2^9 = 512$.

- Action space A: An action is taken at each time step, $a_t \in A$. The action space, A, consists of both the information gathering actions (i.e., which cue to fixate) and decision making actions (i.e., Block/Allow transaction). Therefore, the size of the action space is 11 (9 cues plus 2 decision actions).

- Reward function $r(S,A)$: At any moment, the environment (in one of the states s) generates a reward (cost if the value is negative), $r(s,a)$, in response to the action taken a. For the information gathering actions, the reward is the time cost (the unit is seconds). The time cost includes both the dwell time on the cues and the saccadic time cost of traveling between cues. More detail regarding the time cost is provided in subsection 'Time Cost' below. In the experiment to be modeled (described below), the participants were asked to complete 100 *correct* trials as quickly as possible, so that errors were operationalized as time cost. In the model, the cost for incorrect decisions is based on participants' average time cost (Seconds) for a trial (CT:17 ± 15; CC:24 ± 7;VT:20 ± 8; VC:13 ± 3). That is, the penalty of an incorrect trial is the time cost for doing another trial.

- Transition function $T(S_{t+1}|S_t,A_t)$: In addition to the reward, another consequence of the action is that the environment switches to a new state according to the transition function. In the current task the states (i.e. displayed information patterns) stay unchanged across time steps within one trial. Therefore, $T(S_{t+1}|S_t,A_t)$ equals to 1 only when $S_{t+1} = S_t$. $T(S_{t+1}|S_t,A_t)$ equals 0 otherwise. That is, the state transition matrix is the identity matrix.

- Observation space O: After transitioning to a new state, a new observation is received. The observation, $o_t \in O$, is defined as the information gathered at the time step t. An observation is a 9-element vector, each element of which represents the information gathered for one of the cues. Each element of the observation has three levels, F (fraudulent), N (normal) and U (unknown). For example, one observation

might be represented as [F,N,U,F,U,U,U,N,N]. Therefore the upper bound on the observation space is $3^9 = 19683$.

- Observation function $p(O_t|S_t,A_t)$: The observation function describes how states (information patterns) and actions (fixate locations) combine to yield observations. The availability of information about a cue is dependent on the distance between this cue and the fixation location (eccentricity). In addition, it is known that an object's color is more visible in the periphery than the object's text label [23]. In our model, the observation model is based on the acuity functions reported in [23], where the visibility of an object is dependent on, for example, the object size, the object feature (color or text), and the eccentricity. The observational model is explained in more details in subsection 'Observation Model' below.

- Discount rate γ: At each time step t, the model receives the reward $r(s_t,a_t)$. The optimal strategy is the one that maximizes the expected long-term reward: $E[\sum_{t=0}^{\infty}\gamma^t r(s_t,a_t)]$ given the uncertainty defined by the elements above, where γ is a discount rate, $0 \leq \gamma < 1$.

Observation Model

The observation obtained is constrained by a theory of the limits on the human visual system. Our model assumed that the text information was obtained only when it was fixated. The color information was obtained based on the color acuity function reported in [23]. This function was used to determine the availability of the color information for each cue given the distance between the cues and the fixated location (called *eccentricity*), and the size of the item. Specifically, on each fixation, the availability of the color information was determined by the probabilities defined in Equation (1).

$$P(available) = P(size + X > threshold) \tag{1}$$

where *size* is the object size in terms of visual angle in degrees; $X \sim \mathcal{N}(size, v \times size)$; $threshold = a \times e^2 + b \times e + c$; e is eccentricity in terms of visual angle in degrees. In the model, the function were set with parameter values of v=0.7, b=0.1, c=0.1, a=0.035 as in [23].

Time Cost

The time cost of the information gathering actions consisted of two factors. One is the saccade duration: the time for traveling between cues. The other is the dwell duration: the time for extracting information from cues.

The saccade duration D (in milliseconds) was determined with the following equation from [2]:

$$D = 37 + 2.7A \tag{2}$$

where A is the amplitude (in terms of visual angle in degrees) of the saccade between two successive fixations.

The dwell durations used in the model were determined from experimental data (see below).

Bayesian belief update

At each time step, the environment is in a state, which is not directly observed. The model maintains a belief b about the state given the information observed so far. Every time the agent takes an action a and observes o, b is updated by Bayes' rule. Below we describe how the belief update is computed. An example update from $t = 0$ to $t = 1$ is given, which could be generalized to the update from t to $t + 1$.

The belief is initially assumed to be an uniform distribution over all possible states, b_0. That is, without any evidence, the model believes that the environment is equally possible to be in one of the states. At $t = 1$, an action, a_1, is taken, which causes the environment to transition from state s to state s' with probability $T(s' \mid s,a)$ (the transition function). After reaching s', one observation,o_1, is received with probability $p(o_1 \mid s', a_1)$ (the observation function). The belief state, b_1, is obtained given the action a_1, the observation o_1, and the previous belief b_0, as in Equation (3) below.

$$b_1(s') = \frac{\sum_{s\in S} b_0(s) \times T(s' \mid s,a_1) \times p(o_1 \mid s', a_1)}{\sum_{s'\in S} p(o_1 \mid s', a_1) \sum_{s\in S} T(s' \mid s,a_1) b_0(s)} \tag{3}$$

As mentioned above $T(s'|s,a_1) = 1$ only if $s' = s$, and 0 otherwise, Equation (3) can be simplified as Equation (4):

$$b_1(s) = \frac{b_0(s) \times p(o_1 \mid s,a_1)}{\sum_{s\in S} p(o_1 \mid s,a_1) b_0(s)} \tag{4}$$

At each time t, a belief b_t vector consists of a probability for each of the states, $b_t(s_i)$, where $i \in 1,2,3,...N_s$ and N_s is the number of the states. In our model, the probabilities of the states were rounded to two decimal places. b_t was then used as a prior for next update when a_{t+1} and o_{t+1} is receiving (Equation (5)).

$$b_{t+1}(s_i) = \frac{b_t(s_i) \times p(o_{t+1} \mid s_i, a_{t+1})}{\sum_{s_i\in S} p(o_{t+1} \mid s_i, a_{t+1}) b_t(s_i)} \tag{5}$$

Learning

The control knowledge is represented as a mapping between the beliefs and actions, which is learned with Q-learning [47]. Further details of the algorithm can be found in any standard Machine Learning text (e.g.[47, 41]). Here, a summarized procedure for the computer simulation is provided in box Algorithm 1.

Before learning, an empty Q-table was assumed in which the values (i.e., Q-values) of all belief-action pairs were zero. The model therefore started with no control knowledge and action selection was entirely random. The model was then trained through simulated experience until performance plateaued (requiring about $5 * 10^6$ trials). The model explored the action space using an ε-greedy exploration based on the Q-table. This means that it exploited the greedy/best action with a probability $1 - \varepsilon$, and it explored all the actions randomly with probability ε. ε was set to 0.05 in our model. Q-values of the encountered belief-action pairs were adjusted according to the reward and cost feedback, as shown in the equations in line 10 and line 13. Q(b,a) is the Q-value for one belief-action pair (b,a); r is the immediate reward/cost obtained while the action a is taken (based on the reward function defined); α is called learning rate, which is set to 0.1; γ is called discounted factor,

Algorithm 1 Pseudocode for the simulation

```
Initialize Q(B,A) = 0;
for each simulating trial
    Initialize b to be uniform across all the states
    terminal=FAULT;
    while terminal= FAULT do
        choose a from b (using ε-greedy).
        Take action a, observe o, receive r.
        if a ≠ BLOCK and a ≠ ALLOW then
            Belief update b'
            Update Q(b,a):
Q(b,a) ← Q(b,a) + α[r + γ max_a' (Q(b',a') − Q(b,a))]
            b ← b'
        else
            Update Q(b,a):
            Q(b,a) ← Q(b,a) + α[r − Q(b,a))]
            terminal=TRUE
        end if
    end while
end for
```

which is set to 0.9. The idea is that, these Q-values are learned (or estimated) by simulated experience of the interaction tasks. The true Q-values are estimated by the sampled points encountered during the simulations. The optimal policy acquired through this training was then used to generate the predictions described below (last 1000 trials of the simulation)[1]

EXPERIMENT: CREDIT CARD FRAUD DETECTION

Subjects

Sixteen participants (age: 31.6 ± 7.5; 11 males) from staff and students at University of Birmingham voluntarily participated in the experiment. No financial or other incentives were given. Participants were equally and randomly assigned to four experimental groups which are detailed below.

Apparatus

A customized user interface was created in Matlab (The MathWorks). The participants were seated in front of a 22 inch screen instrumented with an eye tracker (X2-60, Tobii, Sweden) recording gaze data at 60 Hz. The eye tracker was operated through Matlab using the Tobii SDK and Matlab binding.

Design

The participants were asked to take on the role of a credit card fraud analyst at a bank. The task was to decide whether a given transaction should be blocked (prevented from being authorized) or allowed. As shown in each panel of Figure 2, nine information sources were laid out in a 3×3 grid. An operation panel was presented on the right side of the interface, including Block/Allow decision buttons and a feedback window.

The nine cues, as shown in Table 1, were selected as relevant to the detection of credit card fraud based on the literature and discussions with domain experts from FICO, FeedZai and the UK Cards Association. Each cue was presented in binary terms based on rules for fraudulent and non-fraudulent behavior (as shown in Table 1); rules were provided to the participants. The cues had *validities* [0.85, 0.70, 0.65, 0.60, 0.60, 0.60, 0.55, 0.55 and 0.55], where validity was defined as the probability that the cue indicated fraud given that the ground truth of the transaction is fraudulent. Validities were arbitrarily assigned to the nine cues and reflected the observation that high quality cues are relatively rare in many tasks. The location of each cue on the interface was assigned randomly for each participant and stayed constant across all trials. Participants were asked to complete 100 correct trials as quickly as possible. As trials in which an error was made (e.g. blocking a non-fraudulent transaction) did not reduce the total number of correct trials required, errors resulted in time costs.

Cue	Normal	Fraudulent	Validity
#1: Transaction Amount	≤ 500	> 510	0.60
#2: Transaction History	3 small amounts in a row	N/A	0.70
#3: Card Present	YES	NO	0.65
#4: CVV Entered	YES	NO	0.55
#5: Card Issued Bank*	Hanford	NorthWest	0.60
#6: Purchase Made in*	Europe	USA	0.85
#7: Card Expiry Check	≥ 5 days	≤ 4 days	0.55
#8: Transaction Time	6:00-20:00	20:00-6:00	0.60
#9: Type of Goods*	Travel agent	Electrical goods	0.55

Table 1. Information used for the 9 cues. Starred cues (*) were counterbalanced across participants.

The experiment was a $3 \times 2 \times 2$ design. The independent factors were *validity*, *format* and *availability*. *Validity* had 9 levels (grouped into high, medium, and low levels of validity). *Format* had two levels: text vs. color. *Availability* had two levels: visible vs. covered. Format and availability give four user interfaces 2.

- Covered-Text (CT) condition (Figure 2a): The cue information was presented in covered text. In order to check each cue, the participants had to click on the associated button on each cue and wait for 1.5 seconds while a blank screen was shown.
- Covered-Color (CC) condition (Figure 2b): The cue information was presented by color (green for possibly normal, red for possibly fraudulent). As with CT, the information was covered until clicked.
- Visible-Text (VT) condition (Figure 2c): The cue information was presented in text. The information was visible immediately (no mouse-click was required).
- Visible-Color (VC) condition (Figure 2d): The cue information was presented in color and no mouse-click was required to reveal it.

Procedure

Participation in the study started with reading and signing the information sheet and consent form, followed by reading the detailed written study instructions. After eye-tracker calibration, the participants then worked through the first transaction with the opportunity to ask any clarifying questions. After this, participants worked through transactions at their own pace.

The workflow for evaluating a transaction was as follows. On each trial, participants were told to use as much information as

[1]The model was implemented in Matlab and can be downloaded on request from the first author (Email: xxx.xxx@gmail.com).

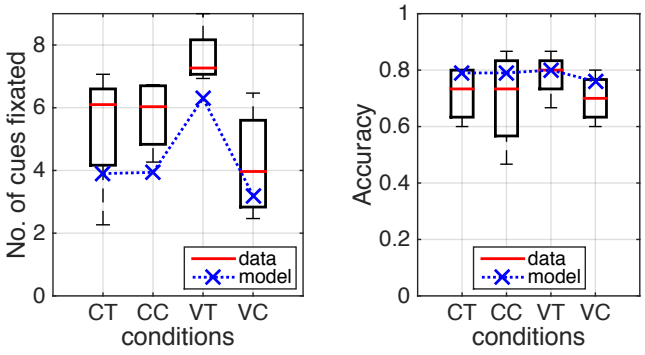

Figure 3. The number of cues (left) and accuracy (right) predicted by the model across 4 experimental conditions (x-axis). Model predictions (blue crosses) are plotted with the participant's data (boxplots). The figure shows that the model predicts that an elevated number of cues will be used by participants in the Visible-Text condition.

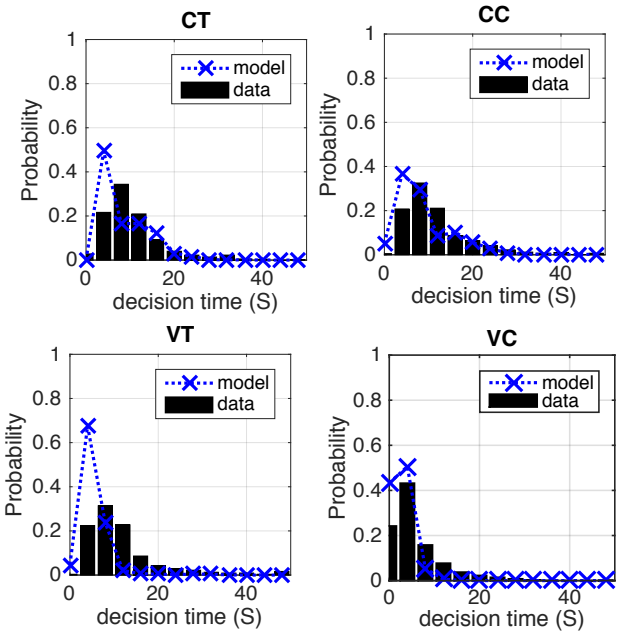

Figure 4. The decision time distributions. The bin centers for the histograms are 0,4,8,12,...48. For human data, the decision time is the time between clicking the 'next transaction' button to clicking 'Allow/Block' decision button. For the model, the decision time includes the saccade time across panels (including information panels and operational panels) and the dwell duration on the panels as described in the 'Time Cost' section.

they see fit (up to 9) by clicking the 'reveal' buttons or by fixating depending on the condition they were assigned to. They then indicated their decision in the right panel by clicking the 'Allow' or 'Block' button. Following this, participants clicked on the cues that they based their decision on (for later cross-validation not reported here) and then confirmed their decision by pressing a 'submit' button. Participants then received feedback regarding the correctness of the decision. The next trial was shown after clicking the button 'next transaction'. This button was placed in the top right corner of the display above the operational panel so as not to confound initial gaze data on the cues. At intervals of 15 minutes, participants were offered the opportunity to take a short break.

RESULTS AND DISCUSSION

On average participants made 43.88 ± 15.68 errors on their way to completing 100 correct trials. In accordance with our theory, we were interested in participant performance once they were adapted to the task. For this reason we analyzed the last 15 trials completed by each participant. The analyzed trials included both correct and error trials. The model results reported below were generated by modeling each individual participant and aggregating.

Information cues fixated and accuracy

As we stated above, the model's dwell times were calibrated to the empirical data for each cue. On average the calibrated dwell time was 0.66 ± 0.10 seconds across the cues.

Having calibrated the model, we were then able to predict how many cues should be fixated in each condition. This prediction and the human data are shown in Figure 3. The model correctly predicts that participants should fixate on more cues in the Visible-Text condition (VT: 6.21 ± 1.32) than in the other three conditions. It also correctly predicts that participants should fixate fewer cues in the Visible-Color condition (Visible-Color: 3.08 ± 1.32) and it learns to be relatively economical in the 'covered' conditions (Covered-Text: 3.95 ± 1.97; Covered-Color: 3.84 ± 1.21).

Intuitively, these findings can be explained in terms of adaptation to information cost and limits of peripheral vision. In the Visible-Text condition, information is cheap and foveated vision is required to read text, therefore more cues are used. In contrast, in the covered conditions (Covered-Text and Covered-Color) information access is expensive reducing the number of cues used. Lastly, in the Visible-Color condition, peripheral vision, rather than only foveated vision, can be used to access information and it appears that as a consequence fewer cues are used, at least by being directly fixated.

A visual comparison of the human data and model predictions (Figure 3) shows that the model accurately predicts the qualitative pattern of results, although the model predicts fewer cues on average (about 1.5 cues fewer).

While participants fixated on more cues in the Visible-Text condition, they did so without significant increase in accuracy (boxplots in the right panel of Figure 3). Participants achieved about 74% accuracy (CT:0.75, CC:0.70, VT:0.78, VC:0.75), which our analysis suggests is close to optimal given the cost of action, the validity of the cues and task of completing 100 tasks as quickly as possible. The model achieved about 79% accuracy across the four conditions (CT:0.79 ± 0.03, CC:0.79 ± 0.01, VT:0.79 ± 0.02, VC:0.76 ± 0.05). The model predicts the same level of accuracy achieved in all conditions (Figure 3).

The variance explained by the model's prediction of the number of cues fixated was $R^2 = 0.72$ with $RMSE = 1.64$. The variance explained by the accuracy prediction was $R^2 = 0.71$ with $RMSE = 0.05$.

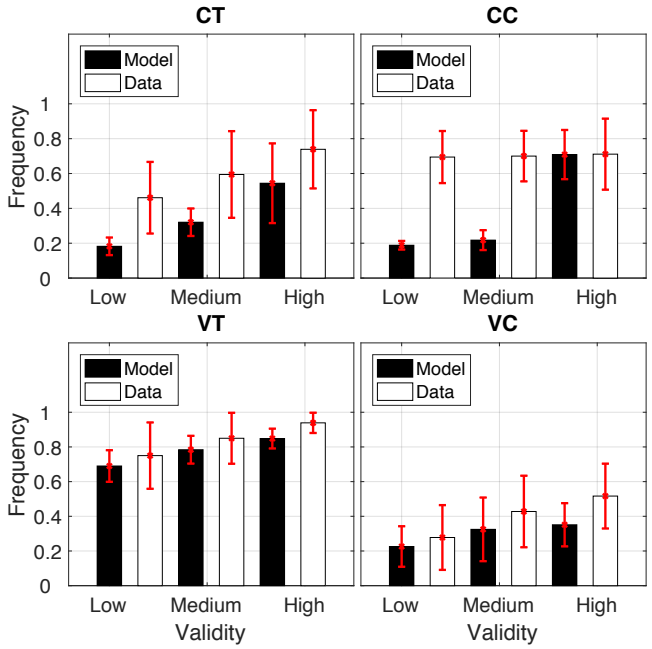

Figure 5. The frequency that the model used cues split by condition and validity. For human data, 'Frequency' (y-axis) represents the frequency that each cue were used by participants in the last 15 trials.

Decision time distribution

It is well known that long-tail left-skewed distributions are a signature of human decision duration (e.g. a review in [10, 5]). Changes in difficulty of decisions produce changes in their spread but very little change in their shape [33, 34]. Figure 4 (black bars) shows that human decision time in our experiment is also long-tailed and left-skewed. The decision-time distributions for the model (Figure 4) show the same shape, although faster than the human data. The distributions are interesting because despite the fact that no skewed distributions are built into the model, they emerge as a consequence of adaptation to the constraints.

The variance explained by the model's decision time prediction for each of the four conditions was (CC) $R^2 = 0.73$, $RMSE = 0.06$; (CT) $R^2 = 0.53$, $RMSE = 0.09$; (VC) $R^2 = 0.87$, $RMSE = 0.07$; (VT) $R^2 = 0.41$, $RMSE = 0.14$. The low R^2 values for the text conditions, CT and VT, are a consequence of the overly fast model performance, which gives a more skewed distribution than for humans.

Effects on frequency of cue fixation

A $2 \times 2 \times 3$ mixed effects ANOVA was used to test the effect of *format* (text vs. color), *availability* (visible vs. covered), and *validity* (low, medium, high) on the frequency of cues fixated. So as to conduct this analysis we first grouped the 9 levels of validity into 3 levels. These were: high validity ([0.85,0.7,0.65], medium validity [0.6, 0.6, 0.6], and low validity [0.55,0.55,0.55]. This grouping reflects the planned structure of the materials but was also necessary to ensure sufficient data in each cell for the analysis (an analysis across all 9 levels proved inconclusive due to high variance). There was

Availability/Format/Validity	Mean ± Std (Data)	Mean ± Std (Model)
covered/text/low	0.46 ± 0.22	0.18 ± 0.03
covered/text/medium	0.59 ± 0.31	0.32 ± 0.04
covered/text/high	0.74 ± 0.24	0.54 ± 0.04
covered/color/low	0.69± 0.10	0.19 ± 0.01
covered/color/ medium	0.70± 0.04	0.22 ± 0.06
covered/color/high	0.71± 0.26	0.71 ± 0.02
visible/text/low	0.75 ± 0.19	0.70 ± 0.12
visible/text/ medium	0.85 ± 0.16	0.78 ± 0.10
visible/text/high	0.94 ± 0.05	0.85 ± 0.05
visible/color/low	0.28 ± 0.24	0.23 ± 0.11
visible/color/medium	0.43 ± 0.24	0.33 ± 0.12
visible/color/high	0.52 ± 0.07	0.35 ± 0.06

Table 2. Cue fixation frequency in each condition for both Data and Model (also see Figure 5)

a significant main effect of validity F(2,24)=4.91, p=0.016, $\eta^2 = 0.29$. There was a significant main effect of format F(1,12)=5.32, p=0.04, η^2=0.31. There was no main effect of availability F(1,12)=0.10, p=0.76, $\eta^2 = 0.008$. There was an interaction between format and availability F(1,12)=13.94, p=0.003, $\eta^2 = 0.54$. There were no other significant interactions.

We also compared the human cue use frequencies to those of the model (Figure 5 and Table 2). The model was not fitted to these data but, rather, provided zero-free parameter predictions given the earlier calibration to dwell times. It is clear from the figure that although the model consistently uses fewer cues than the participants, it does capture the trend for higher frequency use of high validity cues. For the model, this was true irrespective of condition, whereas for the humans it was only true in conditions CT, VT and VC. With the covered-color (CC) interface, participants exhibited a high use of cues irrespective of validity, suggesting, perhaps, that this condition was hardest to learn.

In addition, to capturing the effect of validity, it can also be seen that the model partially captures the effect of format as text cues were fixated more frequently than color cues in both participants and model; VT (highest frequency), CT (next highest) and VC (lowest frequency).

Implications for decision strategies

The analysis above suggests that participants were strategically adapting their use of cues both to the validity of the cues and simultaneously to the structure of the interface and visualisation. Participants chose to use more high validity cues and chose to fixate on more textual cues than color cues, possibly because the color cues could be perceived with peripheral vision. The use of peripheral cues is most evident in that the lowest frequency of fixation on cues was in the Visible Colour condition.

The analysis also provides evidence about the use of decision heuristics such as Take-the Best (TTB) and Weighted Additive (WADD). TTB would be indicated by the participants selecting just the very best cue (which in our interface always discriminates) and then making a Block/Allow decision. However, it can be seen Figure 5 that participants did not only use the highest validity cue. Further, WADD would be indicated by the participants using all of the available cues. However, it

is clear that this is not happening, with participants preferring to use only a subset of the best cues. While there is plenty of evidence in other tasks that people use TTB, they do not seem to do so in our task, and may not do so exclusively.

DISCUSSION

We have reported a new theory of how people use interaction to make decisions in which strategies for information gathering and choice are an emergent consequence of the capacity to make partial observations of a display. This theory can be written precisely by formulating the decision problem as a POMDP for interaction and solving this problem with machine learning to find an optimal decision strategy. Further, we applied the model to different types of information visualization and showed that when color-map visualizations were made available, more use should be made of peripheral vision to gather information. This result is a consequence of the fact that strategies for information gathering can take advantage of the different human acuity functions associated with color and with text; whereas color information can be obtained from the periphery, text must be foveated to be understood.

In addition, the model predicts that people will not make exclusive use of TTB (or WADD). These strategies have been extensively studied in the human decision making literature, but more recent work has suggested that people exhibit a more flexible range of strategies. Instead of assuming TTB or WADD, our model derived the optimal strategy given a POMDP problem formulation; this optimal strategy involved using a weighted integration of the best cues. These cues provide information that optimizes the trade-off between time and accuracy imposed in the experiment. This result is consistent with work that emphasises the adaptation of strategies to a cognitive architecture in interaction with a local task [17, 19, 31]. Further work is required to determine whether TTB emerges as a consequence of different task scenarios.

Also, while not explored in the empirical analysis above, the model makes predictions that are consistent with previously reported strategies. One of the most well-known strategies is called *center-of-gravity* (also called *averaging saccades* or *the global effect*) [12, 46, 44, 45], which refers to the fact that people frequently land saccades on a region of low-interest that is surrounded by multiple regions of high-interest. Figure 6 shows that this 'center-of-gravity' effect is an emergent effect of our model.

The fact that the model is able to predict the *strategies* that people use is a departure from models that are programmed with strategies so as to fit performance time. This is important because it suggests that the theory might be easily developed in the future so as to rapidly evaluate the usability of a broader range of visualizations for a more extensive range of decision tasks. For example, in the near future we wish to consider multidimensional visualizations that not only make use of color, but also size, shape, grouping etc. It should be possible to increment the observation functions, for example with a shape detection capacity, and then use the learning algorithm to find new strategies for the new visualizations.

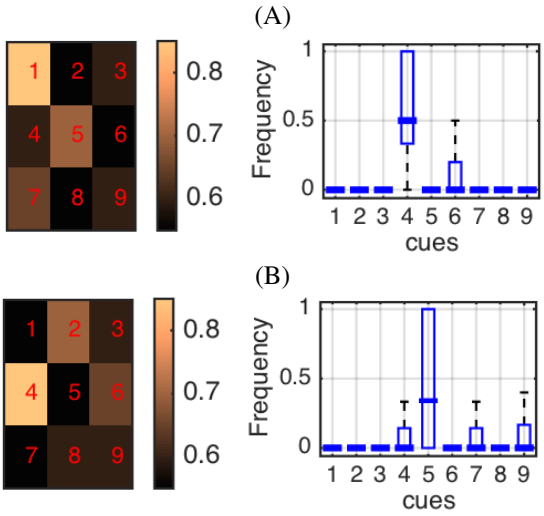

Figure 6. In each row of the figure, the frequency of the model's cue fixations (right panel) is shown for a different spatial arrangement of cue validities (left panel). The validity of ROIs in the left panels is represented as a heat map (high validity is a lighter color). The frequency of model fixations is represented as a box plot. The column numbers in the box plot correspond to the numbers of the ROIs (1..9). In the top row, ROI number 4 has a low validity but is surrounded by relatively high validity ROIs (1,5 and 7). In contrast, in the bottom row, ROI number 5 has a low validity and surrounding ROIs 2, 4 and 6 have high validity. In both rows, the model fixates frequently on the ROI that is surrounded by high validity ROIs. This is known as a centre-of-gravity effect.

Lastly, while we have chosen not to state design implications of our work, we anticipate that there will in the long-term, be design implications. We anticipate that these will derive from future work that is informed by the deeper understanding of interactive decision making that we have provided. This belief is founded on the abundant evidence of the value of theoretical work that can be found in the HCI literature. In fact, some of the field's most influential work has also been some of its most theoretical; see the work of Pirolli and Card on information foraging or the work of Russell et al. (1993) on sensemaking [37]. More generally, we note the potential harm that a focus on design implications may have if it leads to the exclusion of theoretical work [11]. We argue that staying at the cutting edge of interaction theory will help provide a means to improve the efficiency of more applied work by explaining the underlying processes involved in interaction. It is possible that researchers and practitioners who understand the causal structure of interaction are more likely to generate genuinely novel design interventions.

In conclusion, we have reported the first model of how people make decisions through interaction in which strategies for information gathering and decision are derived given a specification of a Partially Observable Markov Decision Process with human-like perceptual observation functions. The model makes predictions that are validated by human data.

ACKNOWLEDGMENTS

This research was supported by the EU-funded SPEEDD project (FP7-ICT 619435).

REFERENCES

1. Chris Baber, Simon Attfield, Gareth Conway, Chris Rooney, and Neesha Kodagoda. 2016. Collaborative sense-making during simulated Intelligence Analysis exercises. *International Journal of Human-Computer Studies* 86 (2016), 94–108.
2. Robert W Baloh, Andrew W Sills, Warren E Kumley, and Vicente Honrubia. 1975. Quantitative measurement of saccade amplitude, duration, and velocity. *Neurology* 25, 11 (1975), 1065.
3. Arndt Bröder and Wolfgang Gaissmaier. 2011. *Heuristics: The foundations of adaptive behavior*. Oxford University Press.
4. Arndt Bröder and Stefanie Schiffer. 2006. Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32, 4 (2006), 904.
5. Jerome R Busemeyer and James T Townsend. 1993. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychological review* 100, 3 (1993), 432.
6. Nicholas J. Butko and Javier R. Movellan. 2008. I-POMDP: An infomax model of eye movement. In *2008 IEEE 7th International Conference on Development and Learning, ICDL*. 139–144.
7. Xiuli Chen, Gilles Bailly, Duncan P. Brumby, Antti Oulasvirta, and Andrew Howes. 2015. The Emergence of Interactive Behavior: A Model of Rational Menu Search. *Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems* 1 (2015), 4217–4226.
8. Andrea Dal Pozzolo, Olivier Caelen, Yann-Aël Le Borgne, Serge Waterschoot, and Gianluca Bontempi. 2014. Learned lessons in credit card fraud detection from a practitioner perspective. *Expert systems with applications* 41, 10 (2014), 4915–4928.
9. Thomas H. Davenport. 2013. How P&G Presents Data to Decision-Makers. *Harvard Business Review* (2013).
10. Fermin del Prado Martin. 2008. A theory of reaction time distributions. (dec 2008).
11. Paul Dourish. 2006. Implications for design. *SIGCHI Conference on Human Factors in Computing Systems (CHI'06)* (2006), 541–550. DOI: http://dx.doi.org/10.1145/1124772.1124855
12. John M. Findlay. 1982. Global visual processing for saccadic eye movements. *Vision Research* 22, 8 (1982), 1033–1045.
13. Wilson S Geisler. 2011. Contributions of ideal observer theory to vision research. *Vision research* 51, 7 (2011), 771–781.
14. Gerd Gigerenzer and Wolfgang Gaissmaier. 2011. Heuristic decision making. *Annual review of psychology* 62 (2011), 451–482.
15. Gerd Gigerenzer and Daniel G Goldstein. 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychological review* 103, 4 (1996), 650.
16. Gerd Gigerenzer and Peter M Todd. 1999. *Fast and frugal heuristics: The adaptive toolbox*. Oxford University Press.
17. Wayne D Gray, Chris R Sims, Wai-Tat W-T Fu, and Michael J Schoelles. 2006. The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psych Review* 113, 3 (2006), 461.
18. Mary Hayhoe and Dana Ballard. 2014. Modeling Task Control of Eye Movements. *Current Biology* 24, 13 (2014), R622—-R628.
19. Andrew Howes, Geoffrey B Duggan, Kiran Kalidindi, Yuan-Chi Tseng, and Richard L Lewis. 2015. Predicting Short-Term Remembering as Boundedly Optimal Strategy Choice. *Cognitive Science* 40, 5 (2015), 1192–1223.
20. Andrew Howes, Richard L. Lewis, and Alonso Vera. 2009. Rational adaptation under task and processing constraints: implications for testing theories of cognition and action. *Psychological review* 116, 4 (2009), 717–751. DOI: http://dx.doi.org/10.1037/a0017187
21. Sanjeev Jha and J Christopher Westland. 2013. A Descriptive Study of Credit Card Fraud Pattern. *Global Business Review* 14, 3 (2013), 373–384.
22. L Kaelbling, Michael L. Littman, and Ar Cassandra. 1998. Planning and Acting in Partially Observable Stochastic Domains. *Artificial Intelligence* 101, 1-2 (1998), 99–134.
23. David E. Kieras and Anthony J. Hornof. 2014. Towards accurate and practical predictive models of active-vision-based visual search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3875–3884.
24. Michael D Lee and S A Zhang. 2012. Evaluating the coherence of Take-the-best in structured environments. *Judgment and Decision Making* 7, 4 (2012).
25. Richard L. Lewis, Andrew Howes, and Satinder Singh. 2014. Computational rationality: linking mechanism and behavior through bounded utility maximization. *Topics in Cognitive Science* 6, 2 (2014), 279–311.
26. Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602* (2013).
27. Ben R Newell and David R Shanks. 2003. Take the best or look at the rest? Factors influencing" one-reason" decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29, 1 (2003), 53.

28. Ben R Newell, Nicola J Weston, and David R Shanks. 2003. Empirical tests of a fast-and-frugal heuristic: Not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes* 91, 1 (2003), 82–96.
29. Jose Nunez-Varela and Jeremy L. Wyatt. 2013. Models of gaze control for manipulation tasks. *ACM Transactions on Applied Perception (TAP)* 10, 4 (2013), 20.
30. Manoj Pandey. 2010. Operational risk forum: A model for managing online fraud risk using transaction validation. *Journal of Operational Risk* 1 (2010), 49.
31. Stephen J. Payne and Andrew Howes. 2013. Adaptive interaction: A utility maximization approach to understanding human interaction with technology. *Synthesis Lectures on Human-Centered Informatics* 6, 1 (2013), 1–111.
32. Rajesh P N Rao. 2010. Decision making under uncertainty: a neural model based on partially observable markov decision processes. *Frontiers in computational neuroscience* 4, November (2010), 146.
33. Roger Ratcliff. 2002. A diffusion model account of response time and accuracy in a brightness discrimination task: fitting real data and failing to fit fake but plausible data. *Psychonomic bulletin {&} review* 9, 2 (jun 2002), 278–291.
34. Roger Ratcliff and Philip L Smith. 2010. Perceptual discrimination in static and dynamic noise: the temporal relation between perceptual encoding and decision making. *Journal of experimental psychology. General* 139, 1 (feb 2010), 70–94.
35. Jörg Rieskamp and Ulrich Hoffrage. 2008. Inferences under time pressure: How opportunity costs affect strategy selection. *Acta psychologica* 127, 2 (2008), 258–276.
36. Jörg Rieskamp and Philipp E Otto. 2006. SSL: a theory of how people learn to select strategies. *Journal of Experimental Psychology: General* 135, 2 (2006), 207.
37. Daniel M Russell, Mark J Stefik, Peter Pirolli, and Stuart K Card. 1993. The cost structure of sensemaking. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*. ACM, 269–276.
38. Daniel Sánchez, M A Vila, L Cerda, and José-Mar\'\ia Serrano. 2009. Association rules applied to credit card fraud detection. *Expert Systems with Applications* 36, 2 (2009), 3630–3640.
39. Guy Shani, Joelle Pineau, and Robert Kaplow. 2013. A survey of point-based POMDP solvers. *Autonomous Agents and Multi-Agent Systems* 27, 1 (2013), 1–51.
40. Nathan Sprague, Dana Ballard, and Al Robinson. 2007. Modeling embodied visual behaviors. *ACM Transactions on Applied Perception (TAP)* 4, 2 (2007), 11.
41. R.S. Sutton and A.G. Barto. 1998. Reinforcement Learning: An Introduction. *IEEE Transactions on Neural Networks* 9, 5 (1998), 1054–1054.
42. Julia Trommershäuser, Paul W Glimcher, and Karl R Gegenfurtner. 2009. Visual processing, learning and feedback in the primate eye movement system. *Trends in Neurosciences* 32, 11 (2009), 583–590.
43. Yuan-Chi Tseng and Andrew Howes. 2015. The adaptation of visual search to utility, ecology and design. *International Journal of Human-Computer Studies* 80 (2015), 45–55.
44. S Van der Stigchel and T C Nijboer. 2011. The global effect: what determines where the eyes land? *Journal of Eye Movement Research* 4, 2 (2011), 1–13.
45. Dustin Venini, Roger W. Remington, Gernot Horstmann, and Stefanie I. Becker. 2014. Centre-of-gravity fixations in visual search: When looking at nothing helps to find something. *Journal of Ophthalmology* 2014, June (2014).
46. Françoise Vitu. 2008. About the global effect and the critical role of retinal eccentricity: Implications for eye movements in reading. *Journal of Eye Movement Research* 2, 3 (2008), 1–18.
47. C Watkins and Peter Dayan. 1992. Q-Learning. *Machine Learning* 8 (1992), 279–292.