Computational Rationality as a Theory of Interaction

Antti Oulasvirta
Aalto University
Finland

Jussi P.P. Jokinen
University of Jyväskylä
Finland

Andrew Howes
University of Birmingham
UK

ABSTRACT

How do people interact with computers? This fundamental question was asked by Card, Moran, and Newell in 1983 with a proposition to frame it as a question about human cognition – in other words, as a matter of how information is processed in the mind. Recently, the question has been reframed as one of adaptation: how do people adapt their interaction to the limits imposed by cognition, device design, and environment? The paper synthesizes advances toward an answer within the theoretical framework of *computational rationality*. The core assumption is that users act in accordance with what is best for them, given the limits imposed by their cognitive architecture and their experience of the task environment. This theory can be expressed in computational models that explain and predict interaction. The paper reviews the theoretical commitments and emerging applications in HCI, and it concludes by outlining a research agenda for future work.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**; **User models**.

KEYWORDS

Cognitive modeling, computational rationality, interaction, reinforcement learning, adaptation, individual differences

ACM Reference Format:
Antti Oulasvirta, Jussi P.P. Jokinen, and Andrew Howes. 2022. Computational Rationality as a Theory of Interaction. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3491102.3517739

1 INTRODUCTION

This paper contributes to theories of the cognitive basis of human–computer interaction (HCI). It does so through the lens of cognitive science, which has been central to answering questions pertaining to information navigation, multitasking, visualization, and input device design, among other matters. Rather than seek to answer each of these questions separately, HCI has searched for *general* theories of cognition and its environment. Particularly influential have been *cognitive architectures* [3, 19], which have contributed to theories of information foraging [95], dual task performance [15], menu selection [16], distraction [104], and visual search [67]. Such theories, while abstractions, are central to the practical aims of HCI and have made contributions to computational design [38, 87], human factors [126], design practice [96], and design education [84]. Recently, human–AI cooperation has added another area in which there is a need for theories of cognition [25, 46, 53].

The theory presented in this paper has grown out of difficulties experienced by the authors, cognitive scientists by training, in applying cognitive architectures. We repeatedly faced the issue that each model needs the modeler to hypothesize how the task is completed, and to code this knowledge as production rules. In other words, the modeler must specify a "recipe", a rule set that specifies the user's procedural skill. Writing these rules is challenging, in part because users are clever at generating unexpected strategies that are hard to identify. This difficulty stems from the fact that architectures such as EPIC [66] and ACT-R [3] admit a very large space of possible strategies. They are not sufficiently constrained for ascertaining which strategies users will actually choose. Moreover, rule systems are brittle. They must be updated if the design or environment changes, and a different rule set is needed for each type of user, limiting applications for design and intelligent interfaces.

These issues have recently been addressed by a new class of theories in efforts to explain *why* people choose some strategies in interaction and not others [1, 22–24, 56, 59, 61, 63, 91]. Consider, for example, explaining why people make certain text-entry errors rather than others, why they tend to skim rather than read web pages, and why they sometimes require hundreds of eye movements to interpret a visualization but at other times only a few. These new theories, known as *computationally rational* theories, explain how observable interaction is a consequence of *adaptation* to constraints (bounds) imposed by cognition and environment.

This paper provides a novel synthesis of *computational rationality* as a theoretical framework. It brings the familiar idea of cognitive architectures together with the idea of bounded optimality [13, 74, 102, 103], rooted in machine learning. While computational rationality, as defined by Lewis et al. [74], was identified as one of the foundational concepts of interaction [52], it has not been reviewed from an HCI perspective. The key idea of the theory is that interactive behavior emerges as a consequence of a *control policy* that is optimally adapted to subjective preferences and to bounds, where the preferences include perceived gains and costs, among them the costs of error, and the bounds are imposed by both an internal environment (the mind) and an external environment (including a device). Various sources of internal and external bounds are illustrated in Figure 1. The bounds limit human information processing – that is, they limit the space of computable strategies available for interaction. We develop the argument that the control policy that is adapted to these bounds interacts with the external environment not directly but only *via* its own internal, or cognitive, environment. Under this view, design does not directly

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3517739

determine behavior but, rather, modifies the external environment of the user and thereby influences the actions that a rational user will take. Design can change the perceived rewards/costs and the perceived availability of actions. What we call "interaction," then, is an emergent consequence of how people choose to behave when constrained by preferences, bounds, and environment. The view is called "computational rationality" because the control policy used to predict interaction is the "rational" (or optimal) policy within the limits imposed by the computations available to the mind.

As with most HCI analyses, practical applications of computational rationality start by determining *the problem* that the user faces. A computationally rational analysis is unlike architecture-based theories, however, in that it does not proceed through the analyst conducting a hierarchical decomposition of the procedural knowledge requirements for the task. Instead, sequential decision theory provides a rigorous formalism for representing the problem, including goals, and the known bounds. A solution (strategy or policy) for the resulting decision problem can then be approximated via reinforcement learning, which yields verifiable predictions of a user's behavior in that situation. Although machine learning is used to make predictions, the goal of modeling in HCI is to explain how people interact with designed systems embedded in real world environments; a very different goal to the goals of AI research.

A key practical implication for HCI is that the modeler need not – indeed must not – specify how a task should be done; rather, reinforcement learning (RL), or some other means of generating approximately optimal policies, is used to derive a *policy*. Decision

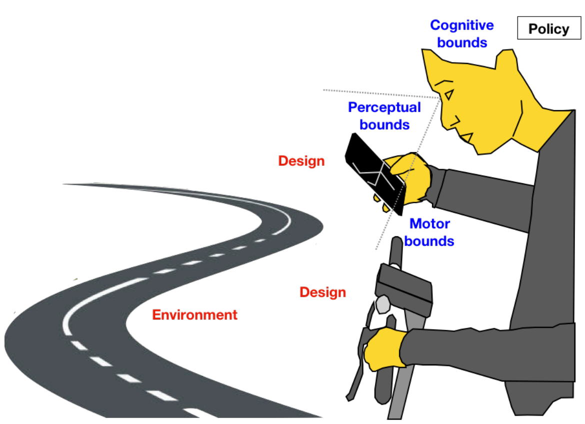


Figure 1: Computationally rational theories assume that users choose behaviors that maximize expected utility, given their bounds. Internal bounds are imposed by cognition and the body and external bounds by the physical environment. Navigating a scooter with a mobile map involves a range of bounds. The phone holder is too low to permit effective use of peripheral vision for collision avoidance, so the rider holds the device with a hand that should be on the handlebars. During attending to the device, the user's uncertainty about the position of the scooter in the lane grows, increasing accident risk. The theory seeks to predict the user's *policy*, or how the user would, in such situation, control locomotion, deploy gaze, and shift the hands.

theoretic formalisms – for example, partially observable Markov decision processes, or POMDPs, which we recruit as a key formalism for this paper – can be used to specify computationally rational theories, and RL methods can be used to solve them. POMDPs provide a general framework for modeling decision processes that are sequential; that is, several steps must be taken to reach a goal or reward, and each step is dependent on the previous state. A key assumption is that the agent cannot directly observe the underlying state of the world. Rather, it must build a representation of the state through observations of small parts of the world (e.g., a single icon) and its actions (e.g., eye movements) – an assumption that is consistent with the HCI setting. Thus, the informal notions of "cognitive strategy" and "heuristic" can be replaced with the formal notion of policy from RL. Moreover, goals can be specified not only as end states but as cumulative reward maximization.

This account is influenced by both cognitive science and machine learning. It draws from cognitive science accounts of the processing limits on what people can do. It takes three things from machine learning: (1) a formal definition of what it means to be computationally rational [74], (2) a way to rigorously specify the user's problem in interaction (the POMDP), and (3) deep RL algorithms for finding solutions to these problems. In contrast to cognitive architectures, computational rationality provides neither a "programming language of the mind" nor scientific hypotheses about the strategies that people use to interact with computers; rather, it makes hypotheses regarding goals and processing limits, after which it predicts strategies, using approximate optimization methods like RL. It therefore avoids the parametric hyper-flexibility that comes with programmed rule systems.

Why should HCI bother with computational rationality? As we argue in the next section, interaction is acutely sensitive to design. User models that assume *fixed* strategies fail to discriminate between designs. Also, awareness of adaptation is a prerequisite for explaining user behavior. An answer to the question "why do people behave like that when given this design?" can be calculated from the person?s goals and bounds. Computational rationality is positioned as a theory focused on the latent causes of user behavior [89], which can be used for inference and prediction.

In summary, this paper makes four contributions toward clarifying computational rationality as a theory of HCI:

(1) It develops the argument that human adaptation is of general interest to the HCI field. We review evidence for this claim (Section 2), and, in Section 3, we provide a synthesis of theoretical assumptions to explain those phenomena.
(2) It provides a unified framework for defining a broad range of computationally rational models, focusing on the problem of integrating cognitive bounds into RL-based models of interaction (Section 4). This approach is contrasted to the standard approach in machine learning.
(3) It provides the first literature review covering model implementations in the area of HCI (Section 5).
(4) It presents an analysis of how computational rationality can help algorithms better make sense of human behavior ("why?") and drive counterfactual reasoning ("what if. . . ?") in HCI, especially in design and adaptation (Section 6).

We end with discussion of a research agenda.

2 ADAPTATION PHENOMENA: WHAT NEEDS TO BE EXPLAINED?

Computational rationality is a theory of interaction as adaptation to bounds, but what are the observable phenomena that constitute adaptation in HCI? To answer this question, we ask *why* people adapt to the types of bounds that are implicated in a wide range of problems faced in computer use. While our list of phenomena is unlikely to be definitive, we have cast the net widely. One overarching theme is related to individuals' differences. People are observed to vary in personality attributes, in preferences, and in cognitive abilities [2, 5, 34, 77, 82, 90, 123]. For example, individual-specific differences are reported in strategies for multimodal interaction, with some people preferring to do one thing at a time (e.g., speak and then point) while others make use of some overlap [90]. Theories of adaptation must be capable of explaining *why* such diversity arises, not just describing how.

2.1 Adaptation to memory limits

Human memory – working memory, semantic memory, episodic memory, etc. – imposes significant bounds on how people adapt to interactive tasks. Much of what we see in the design of interaction is, in part, a response to the properties of human memory. One of the benefits of menus and icons, for example, is that they support recognition memory and do not demand that commands be recalled, the latter being much less reliable. However, humans are not merely victims of limited memory. Instead, people adapt how they perform tasks to be compliant with their particular memory limits. Consider, for example, data-entry tasks, such as copying a phone number from paper to a smartphone. When performing this task, people *choose* to break it up into manageable chunks. Smaller chunks will lead to lower probability of error but take longer since there are additional costs of switching back and forth between paper and phone. People tend to break the task up into chunks of 2, 3, or 4 digits, with most people being capable of using any of these distinct chunk sizes if pressed, but why do some people *prefer* one chunk size over another when given the choice? The answer is that individuals tend to choose a chunk size that makes the best use of their memory, given the task's demands for speed and accuracy [54]. Phenomena such as this must be explained by cognitive theories of HCI.

2.2 Adaptation to perceptual bounds

There are multiple limits imposed by the human visual system that are important in HCI [124]. For example, within 2.5 degrees of eccentricity from the center of the fovea, visual acuity falls by 50% [68]. The user must actively gather information with multiple eye movements, building up a "picture" of the world while doing so. There is evidence that these eye movements are adapted to the perceptual properties of each individual's own pattern of retinal cones and rods [41] as well as to expected information gain [83]. Further, a combination of peripheral vision and foveated vision is known to be crucial to visual search tasks [67], wherein people adapt to the various visual acuity profiles provided by the color, shape, and size of a target. For example, because shape information is the most difficult to perceive in peripheral vision, people are less likely to choose to use it to guide search than color or size. Cognitive theories in HCI must explain how perceptual bounds, such as those of vision, shape interactive behavior.

2.3 Adaptation to motor bounds

Aimed movement is a ubiquitous task in HCI [78]. Because human movements are noisy, an initial aimed movement can end up missing the intended target. This may result in a secondary (corrective) movement or, sometimes, in the user selecting an incorrect button that is adjacent to the target, which can be costly. In a series of studies of ballistic movements toward a flat surface, it was shown that people adapt aimed movement to the amount of motor noise *in their own particular motor system* and to the cost of potential errors [119, 120]. Movements are so noisy that, more often than not, people use multiple submovements in order to make aimed movements. There is evidence suggesting that this multiple-submovement strategy is an optimal adaptation to tasks of non-ballistic aimed movement across a flat surface [80]. Evidence of the planned nature of multiple submovements comes in the observation that people systematically undershoot a target, presumably in an effort to minimize total movement time amid uncertainty [35, 49, 50].

When performing a motor task, a person can choose to do it quickly and less accurately or slowly and more accurately. In one study of speed–accuracy tradeoffs [125], participants were asked to perform a pointing task in conditions that differed in instruction. They were asked to be extremely accurate, accurate, neutral, fast, or extremely fast. Perhaps it is no great surprise that people followed the instructions, but nonetheless a quantitative relationship was observed between speed and accuracy. Others have shown that people can optimize externally imposed speed/accuracy tradeoffs [55, 122].

While HCI researchers are accustomed to thinking about how long it takes a user to point to a target in terms of Fitts's law [78], it does not explain the adaptation of movement time to bounds. Movement times increase as target size decreases, for example, because the effect of perceptual/motor noise is increased, yet Fitts's law does not have a term for noise in its formulation. Users not only take more time as the size and distance of the target increase but fundamentally change how they perform the task. For example, the onset latency of hand and eye movements, along with the interval between them, systematically adapts as the task characteristics change [9]. Such strategies are an adaptive response to noise. This limitation is important for practice. To expand the scope of design decisions beyond target size and width, we need to include strategic adaptations in models of aimed movement.

2.4 Adaptation to the environment

Studies of HCI have revealed a range of ways in which interaction adapts to the environment [51, 106]. For example, to maintain situation awareness while looking at a phone screen, some drivers hold the phone at the top of the steering wheel, some adjust the position of the car in traffic in an effort to enhance safety, and some initiate phone calls at times of low workload (such as at traffic lights) [36]. Adaptations while one is driving also involve interleaving secondary task performance at task boundaries [15]. The point here is that people actively adapt to changes in the environment; if higher uncertainty increases the error rate, then people change the way they perform the task in order to compensate the best they can.

Adaptation to ecology. While HCI researchers often foreground the immediate features of the context, they sometimes forget the shaping effect of the user's *previous experience* [92]. Evidence for this has been instrumental in establishing the "rational" view of human cognition [3, 85]. Human memory is shaped too by the expected distribution of future memory demands [107]. For example, users in HCI are more likely to remember passwords that they are likely to need [39]. Users also leave a web site if the expected benefits of continuing are outweighed by those of going elsewhere, and this tradeoff is tuned to the environmental likelihoods [95]. Further, users adapt the rate at which they enter text into a computer or mobile device to various factors, including the probability of error [97]. One of the elements they adjust is the frequency with which they check for errors in data-entry and thumb-typing tasks that require splitting visual attention across the data, keyboard, and feedback [59, 111]. Users have also been shown to allocate time adaptively across multiple documents, "skimming" the documents so as to find the most valuable information [99]. Similarly, people devote more time to the beginning of a paragraph than to its middle or end, and more time to the beginning of a document than its end. These behaviors suggest that information-seeking strategies adapt to the diminishing returns associated with reading beyond the earlier portions of text [33].

2.5 Summary

Our contention is that all of the diverse phenomena described in this section are explained by a simple principle of adaptation called computational rationality. In the following sections, we describe this theory and explore the progress that has been made in applying it to HCI.

3 THEORETICAL COMMITMENTS

While the theoretical framework has been presented already as a tuple defining an optimization problem [74], here we take a different approach, preferring instead to describe a typical control loop in HCI that conforms to the tuple. Figure 2 illustrates the theoretical framework in terms of a closed-loop control process, presenting a comparison to how this structure is commonly used in AI research where the goal is not to predict human behavior. Our presentation is influenced by recent contributions to cognitive science that emphasize the adaptive nature of human cognition, perception, and motor control [18, 27–29, 47, 48, 56, 75, 76].

In many uses of this formalism (POMDPs) in machine learning, an agent interacts with an external environment. In contrast, in computationally rational models, an *agent* interacts with an *internal* environment via observations and actions, and it interacts with a yoked *external* environment via stimuli and responses. To our knowledge, the first extension of POMDPs to agents with internal environments was proposed by Barto et al. [8] and later Singh and colleagues [110]; however, the first applications in HCI only emerged a decade later (see Section 5). Central to this development has been understanding how to model *human-like bounds* via a POMDP.

In computationally rational theories of interaction, the state of the external environment – which is external to the agent's body – is perceptible via stimuli that result in percepts in the internal environment. The agent observes its own internal environment, including percepts, and is presented with the resulting (partial) observations, along with (subjective) rewards. The agent learns a reward-maximizing policy from experience. Experience consists of repeated episodes of observation, reward, and action. The resulting policy is responsible for determining (internal) actions. Some actions (but not all) lead to responses that change the external state and generate further stimuli.

With this framework, interaction can be explained as the behavioral consequence of a control policy adapted to partial observations and subjective rewards in a contextually specific environment. Underpinning this control process is a set of theoretical commitments, which we elaborate on in the subsections below:

(1) An agent solves bounded optimality problems defined by its internal environment.
(2) The internal environment represents mental states and imposes individually determined bounds on adaptation.
(3) The external environment includes not only a device design but also its spatially and temporally extended context of use (ecology).
(4) Interaction between internal and external environments is itself a source of bounds on adaptation.
(5) Human preferences and goals are represented with a *subjective reward function* that takes as input the internal state of the agent.

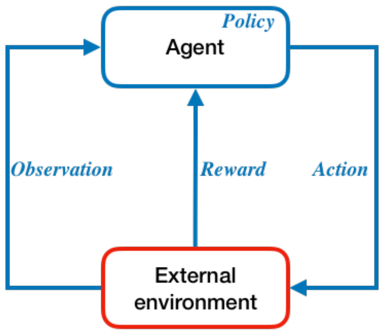
3.1 Agent

An *agent* chooses actions on the basis of its observations. In computational rationality, it does so by solving bounded optimality problems (Commitment 1). The agent adapts by learning a *policy* that is optimal with respect to problems defined through its observations, rewards, and actions. The idea is illustrated in Figure 3.

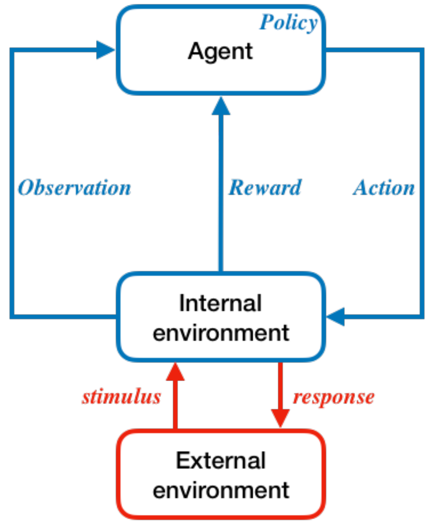
While optimality assumptions are controversial in cognitive science [98], computational rationality is supported by evidence that when one accounts for the bounds imposed on the agent by the internal and external environments, the resulting bounded optimal policy closely models human behavior [6, 17, 23, 41, 43, 48, 56, 57, 64, 69, 79, 85, 116]. A computationally rational theory is a theory of the bounds on adaptation and not the processes by which the policy is achieved.

3.2 Policy

An optimal policy is what determines the selection of actions. It is the agent's solution to the bounded optimization problem. The concept comes from reinforcement learning [115]. In contrast against a cognitive architecture model, the policy is not programmed by a researcher but acquired through experience. Further, a policy is *sequential* in the sense that it is conditioned on the latest observations. It determines how interaction will unfold. A single policy may give rise to a range of strategic behaviors.



(a) A POMDP interacting directly with an external environment



(b) A computationally rational POMDP agent mediated by its internal environment

Figure 2: Computationally rational models can be specified through POMDPs, a formalism for describing sequential decision problems under partial observability; POMDPs explain interactive behavior as a consequence of a policy that is optimally adapted to observations and reward. (a) In machine learning, POMDPs often assume that agents operate directly with external environments. (b) A computationally rational theory uses POMDPs in which the agent does not operate directly with the external environment, doing so only via its internal environment (mental states).

3.3 Internal environment

The internal environment is a theory of the cognitive states on which the agent's policy can be conditioned – as well as the state's dynamics. The primary determinant of the internal environment's state is the percepts constructed from stimuli, but, in addition, the state can include psychological constructs such as memory, emotion, or stress. Capacity limits, among them memory capacities, are an important feature of the internal environment. Moreover, they are individuated, with each person being subject to an individual-specific profile of bounds. This leads to Commitment 2.

In addition, it is assumed that the internal environment is (1) stochastic in that successor states are probabilistically determined. It is also (2) partially observable, because its true state is not available to the agent. Rather, the agent must make repeated (internal) observations and build an estimate of the internal state. For example, the agent may be physically tired but must estimate exactly *how* tired by observing (perhaps repeatedly) its own internal physiology.

3.4 External environment

The external environment refers to the physical context of interaction as well as the interactive technologies contained therein. As the internal environment is, the external environment is stochastic and only partially observable. A sequence of stimuli generated by a user's position with respect to the external environment leads to the construction of percepts in the internal environment. While many computer applications, most computer games among them, are obviously stochastic, also numerous applications that appear deterministic are in fact experienced stochastically. Therefore, both the particulars of a device design and its temporally and spatially extended situation of use contribute to the external environment (Commitment 3).

3.5 Stimulus and response

The intersection between internal and external environments is an important source for bounds (Commitment 4). This is where interaction happens. Both the ability to sense the environmental stimuli and the ability to actuate responses that manipulate the environment involve bounds that are important for explaining interaction. For instance, the human ability to glean visual information from external stimuli is bounded by multiple physiological limitations of the human eye. Only a small foveated portion of the full visual field can be seen accurately, and moving this point of fixation takes time and contains uncertainty in terms of spatial noise. The ability of humans to manipulate the external environment is similarly bounded by noisy response functions. Physical limitations dictate the movement speed and accuracy of a finger, with the human cognition being required to adapt the tradeoff between these to serve a particular goal, given the environment.

3.6 Reward

The reward function is a theory of subjective utility. It encodes whatever is important to the agent: its preferences and goals (Commitment 5) but also the negative rewards associated with states, such as fatigue or time cost. While AI researchers often talk about external environments as generating rewards, our framework commits to the idea that rewards are a function of internal states [8, 110]. External states – for example, the presence of secondary reinforcers – may be associated with rewards but are not rewards in and of themselves. The importance of this separation is that it allows for modeling individual-to-individual differences in the weighting of factors that contribute to reward (e.g., [40]). Weight coefficients, in this sense, define a control point in a class of subjective reward functions that specify the scope of the individual reward functions the model allows.

4 A FORMAL FRAMEWORK FOR THEORIES OF INTERACTION

This section describes a unified framework for building computationally rational models of interaction. It formalizes the concepts introduced in the previous section and illustrated in Figure 2. Our main goal is a flexible and expressive formalization that permits the representation of individual preferences, capacity limits, and designs and is powerful enough to predict interaction when given these bounds. It needs to be expressive enough to account for complex interactive behavior and for individual preferences and capacities that influence how users adapt yet also formal enough to permit implementation in computational models. While formalization of POMDPs is not novel, the goal for this section is to explain how it is used to model HCI where internal bounds are central. The research surveyed in our literature review (in the next section) can be modeled by means of this framework.

4.1 Interaction as sequential decision-making

Following work in cognitive science [27, 29], research on computational rationality in HCI has started to make use of Markovian decision problems. A Markovian decision problem describes a sequential stochastic process in an environment. The environment is in some state, and the probability of it transitioning to another state in the next step depends on the current state and selected action.

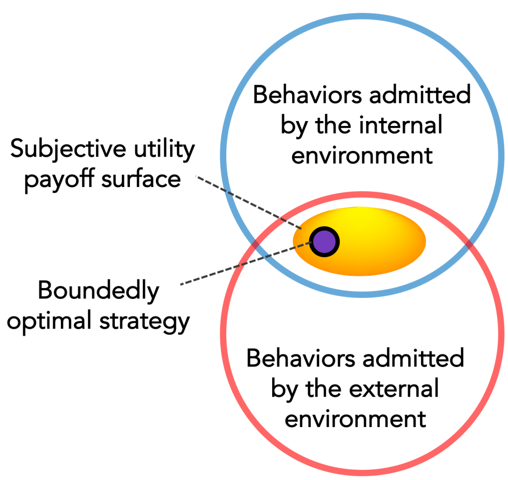


Figure 3: Computationally rational theories make predictions about user behavior by considering how behavior adapts to both external and internal environment. These together determine feasible behaviors, and with RL we can estimate their associated subjective payoffs (rewards vs. costs). Adapted from Howes and Lewis [55].

Here we describe a problem type that is most similar to a partially observable Markov decision problem.

A POMDP is defined as a tuple in terms of the entities introduced in Figure 2. More formally, the POMDP is a tuple $< S, A, T, R, \Omega, O >$ defining a sequential process that in each time step t is in a state $s_t \in S$, where S is the space of possible environmental states. The agent makes an observation $o_{t-1} \in \Omega$, dictated by the observation function O. It takes an action $a_t \in A$, where A is the set of all actions available to the agent. The transition function $T(s_t, a_t, s_{t+1})$ defines the probability $p(s_{t+1} \mid s_t, a_t)$ of the process moving to state s_{t+1} after the action has been performed. A scalar reward $r \in \mathcal{R}$ is then signalled to the agent, as defined by the reward function $R(s_t, a_t) = r$. A key commitment in computational rationality is that the reward is determined by not the external environment but an internal reward-generating process (critic). Depending on the application, the reward function can be defined in terms of the new state (s_{t+1}), an action taken in the current state (s_t, a_t), or the whole triple (s_t, a_t, s_{t+1}). The modeler can choose the states and actions at the level of detail relevant for the current analysis.

The policy determines which actions are chosen, given the most recent observation (see Figure 2). Formally, the decision to take an action in a given state is defined as a policy π, which, for a history of observations or its summary, outputs either a single action $a \in A$ or a probability distribution over all actions A.

4.2 Modeling bounds

Bound is an essential concept in the theory. It refers to anything that constrains the performance of an agent relative to an ideal, unbounded agent. In practice, bounds are modeled by means of the various functions of the POMDP, such as the observation function O and the transition function T. While how best to model bounds remains an open question, we have identified four fundamental types of bounds used so far:

Time refers to a duration, the time it takes to carry out an action or compute something – e.g., system response times and times associated with the user's information-processing capacities. The transition function T is a plausible candidate for modeling such effects (e.g., [61]). *Noise* refers to any external input that obscures the signal of a channel. The central and peripheral nervous system have internal noise. Noise can also have physical origins, caused, for example, by physical contact or the biomechanics of the human body (e.g., [71]). Noise too can be modeled via the transition function T. *Uncertainty* refers to imperfect knowledge – for example, inability to estimate the state of the world or control the outcomes from an action. The cause of uncertainty is often something else, such as partial observability or noise. *Capacity* is a maximum level/amount allowed by some cognitive faculty, such as working memory. Such a capacity limit could be modeled within the transition function T, but other options may exist also.

4.3 Discovering optimal policies

A boundedly optimal agent *does what is best*, given its preferences and the internal environment. The agent can do this because of having found an optimal control policy, which maximizes long-term expected utility, or cumulative rewards. In computational rationality, utility is defined as the mapping from a history of events to a scalar value,

$$\mathcal{U} : \mathcal{H} \rightarrow \mathcal{R}, \tag{1}$$

where $\mathcal{H}$ is the set of all possible histories [74].

Utility is directly connected to the specification of the reward function R when the history space is defined as the space of all possible state-action trajectories of the POMDP. With a given policy π, the utility of an action $a \in A$ in a particular state $s \in S$ is

$$\mathcal{U}_\pi(s, a) = R(s, a) + \gamma \sum_{s' \in S} T(s, a, s')\ \mathcal{U}_\pi(s', a'), \tag{2}$$

where $\gamma \in [0, 1]$ denotes a discount linked to future rewards [101] and $a' \in A$ is the action taken in the next state. The idea is that the utility of an action depends not only on the immediate reward that can be obtained by following said action from this state but also on the future rewards that one can assume follow when the same policy is applied. This is why our scooter driver might find the optimal task-interleaving policy to be a compromise between the two tasks, safe riding and reading the map. The crucial property of utility, visible in Eq. 2, is that it involves self-referential or recursive operations [10]. To know the utility of the present state, one must estimate the utility of future states.

Reinforcement learning is the problem of learning by interacting with the environment [115]. The goal of the learner is to discover an optimal policy π^*. This is a policy that, given Eq. 2, returns the action with the highest expected utility:

$$\mathcal{U}_{\pi^*}(s) = \max_a [R(s, a) + \gamma \sum_{s' \in S} T(s, a, s')\ \mathcal{U}_{\pi^*}(s', a')]. \tag{3}$$

The details of how optimal policies can be discovered are beyond the scope of this paper, since there are multiple RL frameworks and algorithms for solving the problem. While RL has been used to model interactive behavior in prior HCI work (e.g., [72, 127]), the cognitive mechanisms and their bounds have received little attention. Instead, behavior has been studied as an emergent consequence of time costs related to actions. The next section reviews several research papers at the intersection of computational rationality and HCI that utilize RL methods in creating models of human adaptive behavior in interactive tasks.

5 PROGRESS IN MODELING INTERACTIVE BEHAVIOR: A REVIEW

In this section, we provide a review of literature addressing recent progress in developing and validating computationally rational models for HCI tasks. The scope of our literature review is limited to papers that model adaptation in an interactive task as an optimal-policy problem subject to bounds.

5.1 The literature selected

With Google Scholar, we used the following set of search terms: "model AND adapt AND ("computational rationality" OR "rational adaptation" OR "bounded optimality") AND ("interactive task" OR "human-computer interaction" OR "HCI")". The four blocks express our scope: the paper must present a model of adaptation that is based on the idea of optimal adaptation to bounds, and the context must be interaction. Of the 147 results, we excluded master's and PhD theses, public talks, technical and project reports, duplicate entries, workshop papers, and books or book chapters that could not be accessed online, leaving 86 papers. Out of these, 66 were excluded because they did not have an interactive task, did not report on a computational model, were review papers, or did not model human adaptation. A paper was excluded from the final set of papers if it did not include a computational mechanism for deriving optimal policies. Our final list has 15 papers, all of which are described below and listed in Table 1.

The rest of this section is organized into three subsections: the first looks at individual differences, the second memory limits, and the third perceptual-motor bounds.

5.2 The theme of individuals' differences

Computationally rational theories have explained individual-specific differences in a number of task contexts. Jokinen et al. [62] explained differences in risky driving behavior as optimal adaptation to noise in motor control. Noise was hypothesized to affect not only driver actions but also, when they are looking away from the road, drivers' estimates of car position. The authors succeeded in isolating individuals who were at high risk at producing "tail cases": situations of extreme and therefore dangerous lane deviations. Sarcar et al. [105] modeled the effect of individual-specific tremor (signal-independent noise) on touchscreen typing strategies. They used an optimizer to find out how finger movements and proofreading frequency adapt to an individual's tremor level. Combinatorial optimization was then used to find an optimal text-entry design for each individual. Typing-error rates for users with tremor were predicted to fall from 50% to around 5% on a keyboard where letter groups and word predictions were optimized, as compared with a baseline design. In the future, to expand on accounts addressing individual-specific differences, researchers should look at how to model non-neurotypical cases, such as cases of ADHD or autism.

5.3 The theme of adaptation to memory limits

Computationally rational models have been used to explain how user behavior adapts to limits in both short-term (see Subsection 2.1) and long-term memory. In a computationally rational model, different policies are a consequence of different limits to these various systems. Jokinen et al. [63] showed how visual search adapts to limits of visual short-term memory and the availability and accuracy of long-term positional recall. The model explains how details in the UI design, such as element coloring, affect how humans learn UIs and how they adapt to changes in existing designs. This model was successfully used in work that adapted UI design to individuals' history [117].

5.4 The theme of adapting to perceptual-motor bounds

Perceptual and motor bounds can be modeled in the stimulus and observation functions. Chen et al. [23] predicted menu-search strategies as optimal adaptation to visual features available during fixations. They assumed that different visual features – such as shapes

Table 1: Recent efforts at computationally rational modeling of interactive tasks. All papers report on findings in which policies adapt to environmental or cognitive constraints.

Ref.	Task	Paper Title	Finding
[24]	Visual search	A cognitive model of how people make decisions through interaction with visual displays	Visual search policy in decision-making adapts to the visual design of the UI.
[63]	Visual search	Adaptive feature guidance: Modelling visual search with graphical layouts	Visual search policies adapt to the availability of long-term memory information about target features.
[121]	Visual search	The adaptation of visual search to utility, ecology and design	Visual search policies adapt to web site design.
[60]	Multitasking	Modelling drivers' adaptation to assistance systems	A driver's multitasking policy adapts to the presence of a lane assist.
[62]	Multitasking	Bayesian parameter inference for cognitive simulators	Drivers adapt their multitasking policy to individual-specific driving skill.
[14]	Multitasking	Fast or safe? How performance objectives determine modality output choices while interacting on the move	Choice of interaction modality in multitasking adapts to relative task importance.
[59]	Typing	Touchscreen typing as optimal supervisory control	Typing policy adapts to the design of the keyboard.
[105]	Typing	Ability-based optimization of touchscreen interactions	Touchscreen typing policies adapt to the abilities of the user and the design of the keyboard.
[7]	Pointing	The effect of time-based cost of error in target-directed pointing tasks	Pointing policy adapts to the cost of making errors.
[21]	Pointing	Predicting mid-air interaction movements and fatigue using deep reinforcement learning	Mid-air interaction policies adapt to the physical fatigue from the hand movements.
[23]	Menu selection	The emergence of interactive behavior: A model of rational menu search	Visual search in menus adapts to menu design.
[65]	Menu selection	Inferring cognitive models from data using approximate Bayesian computation	Individuals' menu-search policies adapt to long-term knowledge about item positions, and one can infer these parameters by using Bayesian likelihood-free inference.
[73]	Visual decision-making	Informing decisions: How people use online rating information to make choices	Policy for searching of online ratings to aid in decision-making adapts to the ratings' informativeness.
[94]	Decision-making	Probabilistic formulation of the take the best heuristic	Decision-making policy adapts to the availability of efficient heuristics.
[109]	Drawing	Children adapt drawing actions to their own motor variability and to the motivational context of action	Drawing on a touchscreen adapts to motor uncertainty and changes in rewards and penalties.

of menu labels and their lexical contents – can be sampled at different accuracy, depending on their eccentricity (angular distance from the fovea). They showed that, because of this bound, the optimal policy hinges, in a complex way, on the menu's length and organization. Comparing the results against human eye-tracking data, they found a good match.

Jokinen et al. [59] investigated touchscreen typing by hypothesizing an optimal hierarchical control policy governing how the eyes and fingers move across the screen of a cell phone. The bounds included limited visual acuity and motor noise in pointing. Because of uncertainty as to finger position, the eye is needed to guide pointing movements. However, typing errors still occur, necessitating checking what has been typed, during which the eyes cannot effectively guide the fingers. The authors used a computationally rational model to solve the problem of optimal allocation of visual resources between typing and checking, alongside how the pointing movements should balance finger speed and accuracy. The model successfully reproduced several metrics and phenomena observed when humans type, such as inter-key interval, average typing speed, and correlation between typos and proofreading.

Jokinen and Kujala [60], following prior work [61], analyzed how drivers adapt their multitasking policies to the presence of intelligent driving assistance. Replicating human data, a model adapted glancing behavior to external task conditions such as driving speed and the design of the in-car UI. The model assumed that drivers allocate visual attention so as to maximize the joint task utility, given the uncertainty associated with the driving task. From changing the conditions for the task – e.g., adding automatic lane-keeping functionality and varying its reliability – the model predicted that drivers adapt to driving automation by extending their in-car glances, perhaps to the detriment of safe driving [60].

6 WHY AND WHAT IF: EXPLAINING USER BEHAVIOR AND INFORMING DESIGN

Unique to computational rationality is that the explanations take adaptation into account. But how might computational rationality support the practical aims of HCI, such as user research and design? We expand in this section on how computational rationality can play a role in answering the "what if…?"-type and "why?" questions that are so important to design and user research. "Why did the user click this icon?" "What would happen if the system provided no feedback at this stage?" (see [88]). As before, the benefit of computational rationality is that the adaptive responses of users can be taken into account. In this way, it complements supervised learning as an approach to inference and counterfactual reasoning problems in HCI. For example, while supervised learning may be a feasible approach to predicting human interruptibility from labelled sensor data [58], it does require a large amount of data. In contrast, computationally rational models can, in principle, learn such predictions by exploring a simulated problem domain.

In this section, we provide a theoretical overview of explanations and counterfactual scenarios, linking these uses to the theory and modeling formalisms presented in Sections 3 and 4.

6.1 Answering "why?" questions

A model M predicts interactive behaviour given internal parameters θ and external environment parameters ϕ. These parameters express the psychological and physiological bounds that are relevant to the task. Computational rationality asserts that the policy of the model is an optimal policy given θ and ϕ, so we rewrite Eq. 3 in a more general form:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{h\sim<M_{\pi,\theta},\phi>} \mathcal{U}(h), \tag{4}$$

where $h \in \mathcal{H}$ is the history of the interactive episode. The intuition of this formula is that a user is assumed to follow the policy that maximizes the long-term subjective expected utility. With a given history sampled from simulating the parameterized model, a policy that maximizes the utility of the emerging history is used to predict user behavior.

"Why?" questions can be answered by means of parameter inference, determining the most plausible set of parameters θ^* that can be used to describe the user via the model M. This is done by maximizing the likelihood of observed human data y:

$$\theta^* = \arg\max_{\theta} p(y \mid M_\theta, \phi). \tag{5}$$

The most likely set of parameters is therefore those that, for a given model M in an environment ϕ, produce a set of predictions best matching the observed human data. In practice, because they rely on simulation, computationally rational models require a likelihood-free estimation method [45, 65]. The likelihood is computed from the model M of the user interacting with the external environment ϕ. In computational rationality, this policy is assumed to be optimal with regard to the given inferred parameters. A distinct benefit over cognitive architecture models is that possible policies (production systems) do not need to be provided by the researcher. This decreases the dimensionality of the inference problem. An example that uses likelihood-free inference to provide answers to "why?" questions is a computationally rational model of task interleaving [40]. Researchers found that users with (inferred) high discount factors persisted longer in the task before switching to another task. On the other hand, users who tended to discount long-term rewards performed poorly in the task.

6.2 Answering "what if…?" questions

The benefit of a parameterized simulation model is that it can be used to evaluate various counterfactual interventions. We formalize this by specifying that the designer attempts to optimize a set of interventions i^* (often called "designs") that maximize the expected value function $\mathcal{V}$, given the predicted history of behavior sequences that are adaptations to these interventions:

$$i^* = \arg\max_{i} \mathbb{E}_{i,h} \mathcal{V}(h). \tag{6}$$

In a manner similar to the user model's utility function $\mathcal{U}$ (Eq. 1), the value function $\mathcal{V}$ maps a history h to a scalar. Determining $\mathcal{V}$ is problem-dependent, but, for instance, $\mathcal{V}$ can be a function of $\mathcal{U}$, so as to mandate alignment [100]. Alternatively, in addition to encompassing $\mathcal{U}$, $\mathcal{V}$ can include criteria such as a measure of information gain (see, for example, [20]).

The advantage of theorizing on interaction as computational rationality is evident from the formulas presented here. Because humans are adaptive, (eventually) adjusting their policies when the environment changes, the designer can explore the consequences of hypothetical interventions only if it is possible to model and predict how the user will adapt to these changes before they are actually made. There are a few, though not many, reports concerning computationally rational approaches to designing individualized UIs. In a study of text entry, Sarcar et al. [105] used a computationally rational model to design a touchscreen keyboard for users with tremor. They searched keyboard designs and assessed them by predicting WPM and error rates via a model adapted to the relevant design. They found that typing errors can be significantly reduced by grouping several letters on keys and offering word-completion functionality. In another study, a model of visual search and layout learning was used to optimize layouts in line with the user's history (previously seen site designs) [63, 118]. Visual search was modeled as optimal adaptation to perceptual bounds and memory recall for the given site-visit history. The results showed that layouts, personalized through such a model, can significantly speed up visual search.

7 FROM MICRO-HCI TO MACRO-HCI: A RESEARCH AGENDA

Computational rationality is based on a simple but powerful assumption: people do what is best for them, given what they can do. In contrast to programmable cognitive architectures, computational rationality derives policies from given hypotheses about the constraints that bound adaptation. It offers one answer to the call to develop theories of HCI that deal with a "larger number of phenomena at the level of specificity that is needed to inform design decisions" [52, p. 5049]. Advances have been demonstrated in a handful of tasks – with perhaps their strongest manifestations being in typing, multitasking, menu interaction, and decision-making.

However, computational rationality needs to – in Shneiderman's terminology – reach out from micro-HCI, or the realm of user interfaces and interactive tasks, to macro-HCI, or the realm of social experiences and activities [108]. In doing so, it must begin to tackle a broader range of phenomena of human motivations, social experience, and external contexts that are central to understanding HCI in general.

Motivation dynamics. The most critical of these challenges might be *human motivation*: what makes a user pursue some activities and not others? In the models that we have described above, the answer to this question is provided by the reward function, which is a hypothesis about a user's subjective utility. Typically in existing models, the reward, while subjective, is a simple function of some external outcome state, such as navigating to a particular web page, and the costs of arriving at the outcome. In contrast, in computational models of *intrinsic motivation*, such as *curiosity*, states that are novel to the agent are hypothesized to be rewarding in their own right. A curiosity-driven agent can thereby prioritize learning about an environment at the expense of immediate external reinforcers. Other models of intrinsic motivation focus on empowerment, this being the pursuit of perceived control over the states of the environment [44]. We need to reach beyond curiosity, however, to explain why people often seem to engage in behavior that seems unproductive or even detrimental to them, such as leaving smartphone notifications on while trying to focus on work, or why they keep scrolling through news though knowing that this is not good for them. According to self-determination theory [30], behavior is driven by motivations, which evolve over time in an interplay of beliefs, experiences, and basic needs. The user distracted by notifications may hold inaccurate beliefs about notifications' consequences or be unable to estimate the effects of turning them off. Better understanding of motivation-related dynamics is a key to behavior-change and rehabilitation applications. A related challenge involves emotions. Our literature review revealed that current modeling does not consider emotions. However, this is by no means evidence that computational rationality is not conducive to treating humans as emotional beings. Indeed, some recent studies have used RL to model human emotions [81]. If computationally rational models of interaction are to gain traction in HCI, one can expect emotion to be among the phenomena investigated.

Human learning. Computationally rational models of interaction are models of adapted behavior, not models of the process of adaptation. Although RL [115] is a general statement of the problem of learning from the environment and not targeted specifically at explaining how humans learn, there is ample scientific evidence for its biological basis [27]. Also, humans employ various RL systems when adapting, such as model-free, model-based, and episodic [42]. These likely have an important role to play in explaining interaction, but, to our knowledge, no current model in HCI explicitly takes advantage of these ideas. There is a related challenge in the human ability to *generalize* skills. Without understanding this ability, computer applications will underestimate human abilities in novel encounters, such as facing a previously unseen part of a user interface. For example, in contrast to people who can transfer previously learned control solutions, the text-entry model described above learns a policy for single-finger text entry but would have to be retrained for two- or multi-finger text entry. While model-free RL has been shown to be capable of predicting human adaptations in narrow task domains, it can be difficult to extend these models to broader activities without addressing this. Research in cognitive sciences hints that planning, memory, and supervisory control are exploited to overcome this limitation, as in models of human cognition that employ model-based [31], episodic [42], and hierarchical RL [12]. Probabilistic program induction approaches learning of motor programs by inducing them from experience and combining them like computer programs can combine scripts [70]. Future work should look at how generalization is achieved by combining symbolic capabilities (e.g., reasoning from concepts) and subsymbolic ones.

Situations. A key theory for understanding how people interact with computers is that of situated action [114]. Suchman argued that "plans" are not so much mental control structures that universally precede actions as they are resources produced and exploited within activities. Action is determined not only by plans but also by embodied skills conditioned on the particulars of the immediate situation [112, 113]. Indeed, the highly top-down nature of GOMS-like, pre-programmed production-rule models of cognition represents a commitment that is misaligned with empirical data pointing to the embodied and highly reactive nature of cognition. In contrast, computational rationality is aligned well with the concept of embodied action. Boundedly optimal policies are tuned to the specific conditions of each state that lead to the maximization of reward. They are embodied in the sense that the states are internal ones determined by limited perceptual mechanisms, and they are situated in that they are determined by the dynamics of the external environment as it is experienced by the model. For providing a richer account of situations, one challenge for computational rationality is – in line with Suchman's advice – to understand the relationship between plans and action [86, 92]. An essential aspect of this is to understand the internal–external transition function – in other words, the factors that people pay attention to when choosing what to do.

Context. Context is widely believed to be an important determinant of interactive behavior, but what "context" means in the HCI domain is contested [32]. For some, context is construed as a relatively stable "place" in time, space, and society where interaction occurs over some interval (context 1). For example, using a mobile phone on a train is different from using it in an office. For others, context arises spontaneously as a *consequence* of interaction (context 2), rather than as a location where interaction happens. Context might, for example, be determined moment to moment by what is relevant for an unfolding conversation. The context changes constantly throughout a conversation as the situation develops. Both meanings of context present special challenges and limits to computational rationality as a theory of interaction. For context 1, the future looks relatively promising. All that researchers need do is determine generally relevant aspects of context and build these into the specifications of the decision problems. Then, RL would generate the relevant context-specific adaptations. With context 2, the future looks less bright, since it points to infinite-dimensionality interaction and the potential impossibility of anticipating and modeling a reasonable subset of the factors that may determine the socially constructed, continually changing context. The consequent research challenge is to identify and define universal, lower-dimensionality representations of contexts that can capture reasonable parts of everyday interactions.

Distributed cognition. Advocates of distributed cognition have posited that one general principle of interaction is that "people offload cognitive effort to the environment whenever practical" [51, p. 181]. For example, it is assumed that the more a cockpit can do by way of "remembering its speed," the less a pilot will do. An alternative that is consistent with computational rationality, is that people distribute information processing to the environment only to the extent that it allows them to make adaptive use of their internal processing capabilities. In other words, people adaptively distribute cognition [92, 93]. Some advances have been made toward developing computationally rational accounts of these phenomena [54], but much more work is needed to demonstrate the adequacy of computational rationality as an explanation of distributed cognition phenomena.

Social interaction. Social interaction poses a significant challenge not only to computational rationality but to all cognitive modeling. On the one hand, behavior in mediated communication and collaboration is adaptive. When people interact with each other, they adjust the way they perform tasks to the distribution of relationships with the members of the audience [26]. People using social media are known to adopt a range of strategies in efforts to prevent context collapse [11, 26]. On the other hand, it is not clear how adaptation in computer-mediated social contexts should be modeled. One starting point for modeling is social cognition [37]. It characterizes the mechanisms by which people process, store, and apply information about other people and social situations. This could afford a way to model social contexts by means of constructs that are already available, such as perceptions, beliefs, memory, and reasoning.

Computational design. Finally, while computational rationality holds promise for computational design, it faces technical challenges that have limited its application to some fairly simple problems. Most prior research on computational design in HCI has exploited non-adaptive models (e.g., [87]). This limits them to relatively simple sensorimotor tasks that assume no change in the user's strategy between designs. For instance, SUPPLE used Fitts's law and a readability heuristic to generate widget layouts that accommodate users with motor deficiencies [38]. Computational rationality starts with the notion that user behavior is not static but adapts to design. For example, eye-hand coordination strategies might change. Hence, computational rationality could extend the scope of computational design and improve the outcome quality. Today's training times for computationally rational models are prohibitively long for large design spaces, however. The optimizers have to be smarter, policy learning needs to be faster, or both.

Adaptive and cooperative interfaces. Further into the future, computational rationality could offer a way for HCI researchers to contribute to the development of adaptive and cooperative interfaces [25]. From an HCI perspective, cooperative and mixed-initiative systems should try to maximize the added value that automation gives users while considering the costs and attentional dynamics involved. This requires a strong inferential capability [53]. In machine learning, one of the key methods for inference, inverse reinforcement learning, has been shown to be intractable unless there are assumptions made about bounds [4]. Computational rationality offers exactly that: the better the assumptions about utility, bounds, and environment we put into the model, the better its results. That the approach is built around constructs rooted in psychology has further benefits for applications. An AI application employing such constructs may be better able to explain its behavior and communicate with human partners. Moreover, the fact that a computationally rational policy establishes a causal link between bounds and behavior also acts in favor of explanation [56].

8 SUMMARY

Computational rationality is emerging as a new direction for understanding interaction as an adaptive control process. It recruits several key ideas from cognitive science and machine learning for defining adaptation problems and solving them by means of reinforcement-learning algorithms. We have argued that a deviation from the standard account of machine learning is necessary, to open the door for modeling interactive tasks with humans. In particular, cognition and its bounds should be modeled as a user's internal environment. Building on this assumption, modeling has been performed for a broad range of interactive settings in a generative fashion. Much work remains, however, for expanding the scope of the theory from micro- to macro-HCI. This includes efforts to understand how to model human motivations, contexts, learning, and social interactions.

ACKNOWLEDGMENTS

This work was funded by the Finnish Center for AI and Academy of Finland ("BAD" and "Human Automata"). We thank our reviewers, Xiuli Chen, Joerg Mueller, Christian Guckelsberger, Sebastiaan de Peuter, Samuel Kaski, Pierre-Alexandre Murena, Antti Keurulainen, Suyog Chandramouli, and Roderick Murray-Smith for their comments.

REFERENCES

[1] Aditya Acharya, Andrew Howes, Chris Baber, and Tom Marshall. 2018. Automation reliability and decision strategy: A sequential decision making model for automation interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE: Los Angeles, CA, 144–148.
[2] Sander Ackermans, Debargha Dey, Peter Ruijten, Raymond H. Cuijpers, and Bastian Pfleging. 2020. The effects of explicit intention communication, conspicuous sensors, and pedestrian attitude in interactions with automated vehicles. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.*
[3] John R. Anderson. 1990. *The adaptive character of thought.* Psychology Press.
[4] Stuart Armstrong and Sören Mindermann. 2018. Occam's razor is insufficient to infer the preferences of irrational agents. In *Advances in neural information processing systems.* 5598–5609.
[5] Zahra Ashktorab, Mohit Jain, Q. Vera Liao, and Justin D. Weisz. 2019. Resilient chatbots: Repair strategy preferences for conversational breakdowns. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.*
[6] B. Bahrami, K. Olsen, P. E. Latham, A. Roepstorff, G. Rees, and C. D. Frith. 2010. Optimally interacting minds. *Science* 329, 5995 (2010), 1081–1085.

[7] Nikola Banovic, Tovi Grossman, and George Fitzmaurice. 2013. The effect of time-based cost of error in target-directed pointing tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1373–1382.

[8] Andrew G. Barto, Satinder Singh, and Nuttapong Chentanez. 2004. Intrinsically motivated learning of hierarchical collections of skills. In *Proceedings of the 3rd International Conference on Development and Learning*. 112–119.

[9] Harold Bekkering, Jos J. Adam, Herman Kingma, A. Huson, and H. T. A. Whiting. 1994. Reaction time latencies of eye and hand movements in single- and dual-task conditions. *Experimental Brain Research* 97, 3 (1994), 471–476.

[10] Richard Bellman. 1966. Dynamic programming. *Science* 153, 3731 (1966), 34–37.

[11] Jens Binder, Andrew Howes, and Alistair Sutcliffe. 2009. The problem of conflicting social spheres: Effects of network structure on experienced tension in social network sites. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*. ACM, 965–974.

[12] Matthew Botvinick and Ari Weinstein. 2014. Model-based hierarchical reinforcement learning and human action control. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, 1655 (2014), 20130480.

[13] Michael E. Bratman, David J. Israel, and Martha E. Pollack. 1988. Plans and resource-bounded practical reasoning. *Computational Intelligence* 4, 3 (1988), 349–355.

[14] Duncan P. Brumby, Samantha C. E. Davies, Christian P. Janssen, and Justin J. Grace. 2011. Fast or safe? How performance objectives determine modality output choices while interacting on the move. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 473–482.

[15] D. P. Brumby, D. D. Salvucci, and A. Howes. 2007. Dialing while driving? A bounded rational analysis of concurrent multi-task behavior. In *Proceedings of the 8th International Conference on Cognitive Modeling*. 121–126.

[16] Michael D. Byrne. 2001. ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human–Computer Studies* 55, 1 (2001), 41–84.

[17] Frederick Callaway, Rangel Antonio, and Griffith Tom. 2020. Fixation patterns in simple choice are consistent with optimal use of cognitive resources. *PsyArXiv preprint: https://doi.org/10.31234/osf.io/57v6k* (2020).

[18] Frederick Callaway, Antonio Rangel, and Thomas L. Griffiths. 2021. Fixation patterns in simple choice reflect optimal information sampling. *PLoS Computational Biology* 17, 3 (2021), e1008863.

[19] S. K. Card, T. P. Moran, and A. Newell. 1983. *The psychology of human–computer interaction*. Lawrence Erlbaum: Hillsdale, NJ.

[20] Daniel R. Cavagnaro, Richard Gonzalez, Jay I. Myung, and Mark A. Pitt. 2013. Optimal decision stimuli for risky choice experiments: An adaptive approach. *Management Science* 59, 2 (2013), 358–375.

[21] Noshaba Cheema, Laura A. Frey-Law, Kourosh Naderi, Jaakko Lehtinen, Philipp Slusallek, and Perttu Hämäläinen. 2020. Predicting mid-air interaction movements and fatigue using deep reinforcement learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*.

[22] Xiuli Chen, Aditya Acharya, Antti Oulasvirta, and Andrew Howes. 2021. An adaptive model of gaze-based selection. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.

[23] Xiuli Chen, Gilles Bailly, Duncan P. Brumby, Antti Oulasvirta, and Andrew Howes. 2015. The emergence of interactive behaviour: A model of rational menu search. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 4217–4226.

[24] Xiuli Chen, Sandra Dorothee Starke, Chris Baber, and Andrew Howes. 2017. A cognitive model of how people make decisions through interaction with visual displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1205–1216.

[25] Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson, and Thore Graepel. 2021. Cooperative AI: machines must learn to find common ground.

[26] Jenny L. Davis and Nathan Jurgenson. 2014. Context collapse: Theorizing context collusions and collisions. *Information, Communication & Society* 17, 4 (2014), 476–485.

[27] Nathaniel D. Daw. 2014. Advanced reinforcement learning. *Neuroeconomics* (2014), 299–320.

[28] P. Dayan. 2014. Rationalizable irrationalities of choice. *Topics in Cognitive Science* (2014).

[29] Peter Dayan and Nathaniel D. Daw. 2008. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience* 8, 4 (2008), 429–453.

[30] Edward L. Deci and Richard M. Ryan. 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry* 11, 4 (2000), 227–268.

[31] Bradley B. Doll, Dylan A. Simon, and Nathaniel D. Daw. 2012. The ubiquity of model-based reinforcement learning. *Current Opinion in Neurobiology* 22, 6 (2012), 1075–1081.

[32] Paul Dourish. 2004. What we talk about when we talk about context. *Personal and ubiquitous computing* 8, 1 (2004), 19–30.

[33] G. B. Duggan and S. J. Payne. 2009. Text skimming: The process and effectiveness of foraging through text under time pressure. *Journal of Experimental Psychology: Applied* 15, 3 (2009), 228–242.

[34] Dennis E. Egan. 1988. Individual differences in human–computer interaction. In *Handbook of human–computer interaction*. Elsevier, 543–568.

[35] Sascha E. Engelbrecht, Neil E. Berthier, and Laura P. O'Sullivan. 2003. The undershoot bias: Learning to act optimally under uncertainty. *Psychological Science* 14, 3 (2003), 257–261.

[36] Mattias Esbjörnsson, Oskar Juhlin, and Alexandra Weilenmann. 2007. Drivers using mobile phones in traffic: An ethnographic study of interactional adaptation. *International Journal of Human–Computer Interaction* 22, 1–2 (2007), 37–58.

[37] Susan T. Fiske and Shelley E. Taylor. 1991. *Social cognition*. McGraw–Hill.

[38] Krzysztof Gajos and Daniel S. Weld. 2004. SUPPLE: Automatically generating user interfaces. In *Proceedings of the 9th International Conference on Intelligent User Interfaces*. 93–100.

[39] Xianyi Gao, Yulong Yang, Can Liu, Christos Mitropoulos, Janne Lindqvist, and Antti Oulasvirta. 2018. Forgetting of passwords: Ecological theory and data. In *Proceedings of the 27th USENIX Security Symposium*. USENIX, 221–238.

[40] Christoph Gebhardt, Antti Oulasvirta, and Otmar Hilliges. 2021. Hierarchical reinforcement learning explains task interleaving behavior. *Computational Brain & Behavior* 4 (2021), 284–304.

[41] Wilson S. Geisler. 2011. Contributions of ideal observer theory to vision research. *Vision Research* 51, 7 (2011), 771–781.

[42] Samuel J. Gershman and Nathaniel D. Daw. 2017. Reinforcement learning and episodic memory in humans and animals: An integrative framework. *Annual Review of Psychology* 68 (2017), 101–128.

[43] Gerd Gigerenzer. 2018. The bias bias in behavioral economics. *Review of Behavioral Economics* 5, 3-4 (2018), 303–336.

[44] Christian Guckelsberger, Christoph Salge, Jeremy Gow, and Paul Cairns. 2017. Predicting player experience without the player: An exploratory study. In *Proceedings of the Annual Symposium on Computer–Human Interaction in Play*. 305–315.

[45] Michael U. Gutmann and Jukka Corander. 2016. Bayesian optimization for likelihood-free inference of simulator-based statistical models. *Journal of Machine Learning Research* 17, 125 (2016), 1–47.

[46] Dylan Hadfield-Menell, Anca Dragan, Pieter Abbeel, and Stuart Russell. 2016. Cooperative inverse reinforcement learning. *arXiv preprint: 1606.03137* (2016).

[47] Ulrike Hahn. 2014. The Bayesian boom: Good thing or bad? *Frontiers in Psychology* 5 (2014).

[48] U. Hahn and P. A. Warren. 2009. Perceptions of randomness: Why three heads are better than four. *Psychological Review* 116, 2 (2009), 454–461.

[49] Christopher M. Harris. 1995. Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study. *Vision Research* 35, 5 (1995), 691–701.

[50] David B. Henson. 1978. Corrective saccades: Effects of altering visual feedback. *Vision Research* 18, 1 (1978), 63–67.

[51] James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed cognition: Toward a new foundation for human–computer interaction research. *ACM Transactions on Computer–Human Interaction* 7, 2 (2000), 174–196.

[52] Kasper Hornbæk and Antti Oulasvirta. 2017. What is interaction?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 5040–5052.

[53] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 159–166.

[54] Andrew Howes, Geoffrey B. Duggan, Kiran Kalidindi, Yuan-Chi Tseng, and Richard L. Lewis. 2016. Predicting short-term remembering as boundedly optimal strategy choice. *Cognitive Science* 40, 5 (2016), 1192–1223.

[55] Andrew Howes and Richard L. Lewis. 2018. The role of (bounded) optimization in theory testing and prediction. *Behavioral and Brain Sciences* 41 (2018).

[56] 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.

[57] Andrew Howes, Paul A. Warren, George Farmer, Wael El-Deredy, and Richard L. Lewis. 2016. Why contextual preference reversals maximize expected value. *Psychological Review* 123, 4 (2016), 368–391.

[58] Scott Hudson, James Fogarty, Christopher Atkeson, Daniel Avrahami, Jodi Forlizzi, Sara Kiesler, Johnny Lee, and Jie Yang. 2003. Predicting human interruptibility with sensors: A Wizard of Oz feasibility study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 257–264.

[59] Jussi Jokinen, Aditya Acharya, Mohammad Uzair, Xinhui Jiang, and Antti Oulasvirta. 2021. Touchscreen typing as optimal supervisory control. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.

[60] Jussi P. P. Jokinen and Tuomo Kujala. 2021. Modelling drivers' adaptation to assistance systems. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 12–19.

[61] Jussi P. P. Jokinen, Tuomo Kujala, and Antti Oulasvirta. 2020. Multitasking in driving as optimal adaptation under uncertainty. *Human Factors* (2020), 0018720820927687.

[62] Jussi P. P. Jokinen, Ulpu Remes, Tuomo Kujala, and Jukka Corander. 2021. Bayesian parameter inference for cognitive simulators. In *Bayesian methods for interaction design (forthcoming)*, Nikola Banovic, Per Ola Kirstensson, Antti Oulasvirta, and John H. Williamson (Eds.). Cambridge University Press.

[63] Jussi P. P. Jokinen, Zhenxin Wang, Sayan Sarcar, Antti Oulasvirta, and Xiangshi Ren. 2020. Adaptive feature guidance: Modelling visual search with graphical layouts. *International Journal of Human–Computer Studies* 136 (2020), 102376.

[64] Keno Juechems, Jan Balaguer, Bernhard Spitzer, and Christopher Summerfield. 2021. Optimal utility and probability functions for agents with finite computational precision. *Proceedings of the National Academy of Sciences* 118, 2 (2021).

[65] Antti Kangasrääsiö, Jussi P. P. Jokinen, Antti Oulasvirta, Andrew Howes, and Samuel Kaski. 2019. Parameter inference for computational cognitive models with Approximate Bayesian Computation. *Cognitive Science* 43, 6 (2019), e12738.

[66] David E. Kieras. 1997. Building cognitive models with the EPIC architecture for human cognition and performance. *Electrical Engineering* 12 (1997), 48109–48109.

[67] David E. Kieras and Anthony J. Hornof. 2014. Towards accurate and practical predictive models of active-vision-based visual search. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3875–3884.

[68] David E. Kieras and Anthony J. Hornof. 2014. Towards accurate and practical predictive models of active-vision-based visual search. In *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3875–3884.

[69] Konrad P. Körding and Daniel M. Wolpert. 2006. Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences* 10, 7 (2006), 319–326.

[70] Brenden M. Lake, Ruslan Salakhutdinov, and Joshua B. Tenenbaum. 2015. Human-level concept learning through probabilistic program induction. *Science* 350, 6266 (2015), 1332–1338.

[71] Byungjoo Lee and Hyunwoo Bang. 2015. A mouse with two optical sensors that eliminates coordinate disturbance during skilled strokes. *Human–Computer Interaction* 30, 2 (2015), 122–155.

[72] Katri Leino, Antti Oulasvirta, and Mikko Kurimo. 2019. RL-KLM: Automating keystroke-level modeling with reinforcement learning. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 476–480.

[73] Stelios Lelis and Andrew Howes. 2011. Informing decisions: How people use online rating information to make choices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2285–2294.

[74] 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.

[75] Falk Lieder and Thomas L. Griffiths. 2017. Strategy selection as rational metareasoning. *Psychological Review* 124, 6 (2017), 762–794.

[76] Falk Lieder and Thomas L Griffiths. 2019. Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences* (2019), 1–85.

[77] Wendy E. Mackay. 1988. Diversity in the use of electronic mail: A preliminary inquiry. *ACM Transactions on Information Systems (TOIS)* 6, 4 (1988), 380–397.

[78] I. Scott MacKenzie. 2018. Fitts' law. *Handbook of human–computer interaction* 1 (2018), 349–370.

[79] Laurence T. Maloney and Pascal Mamassian. 2009. Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. *Visual Neuroscience* 26, 1 (2009), 147–155.

[80] D. E. Meyer, R. A. Abrams, S. Kornblum, C. E. Wright, and Keith Smith. 1988. Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review* 95, 3 (1988), 340–370.

[81] Thomas M. Moerland, Joost Broekens, and Catholijn M. Jonker. 2018. Emotion in reinforcement learning agents and robots: A survey. *Machine Learning* 107, 2 (2018), 443–480.

[82] Sean A. Munson and Paul Resnick. 2010. Presenting diverse political opinions: How and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1457–1466.

[83] Jiri Najemnik and Wilson S. Geisler. 2008. Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision* 8, 3 (2008), 4.

[84] Donald A. Norman. 1988. *The psychology of everyday things.* Basic Books.

[85] M. Oaksford and N. Chater. 2007. *Bayesian rationality: The probabilistic approach to human reasoning.* Oxford University Press: Oxford, UK.

[86] K. P. O'Hara and S. J. Payne. 1998. The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology* 35, 1 (1998), 34–70. https://doi.org/10.1006/cogp.1997.0676

[87] Antti Oulasvirta, Niraj Ramesh Dayama, Morteza Shiripour, Maximilian John, and Andreas Karrenbauer. 2020. Combinatorial optimization of graphical user interface designs. *Proc. IEEE* 108, 3 (2020), 434–464.

[88] Antti Oulasvirta and Kasper Hornbæk. 2021. Counterfactual thinking: What theories do in design. *International Journal of Human–Computer Interaction* 38, 1 (2021), 78–92.

[89] Antti Oulasvirta, Per Ola Kristensson, Xiaojun Bi, and Andrew Howes. 2018. *Computational interaction.* Oxford University Press.

[90] Sharon Oviatt, Rebecca Lunsford, and Rachel Coulston. 2005. Individual differences in multimodal integration patterns: What are they and why do they exist?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 241–249.

[91] John W. Payne, James R. Bettman, and Eric J. Johnson. 1993. *The adaptive decision maker.* Cambridge University Press.

[92] 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.

[93] S. J. Payne, A. Howes, and W. R. Reader. 2001. Adaptively distributing cognition: A decision-making perspective on human–computer interaction. *Behaviour & Information Technology* 20, 5 (2001), 339–346.

[94] Tomi Peltola, Jussi Jokinen, and Samuel Kaski. 2019. Probabilistic formulation of the Take The Best heuristic. *arXiv preprint: 1911.00572* (2019).

[95] P. Pirolli. 2007. *Information foraging theory: Adaptive interaction with information.* Vol. 2. Oxford University Press: USA.

[96] Peter G. Polson, Clayton Lewis, John Rieman, and Cathleen Wharton. 1992. Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. *International Journal of Man–Machine Studies* 36, 5 (1992), 741–773.

[97] Philip Quinn and Shumin Zhai. 2016. A cost–benefit study of text entry suggestion interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 83–88.

[98] Dobromir Rahnev and Rachel N. Denison. 2018. Suboptimality in perceptual decision making. *Behavioral and Brain Sciences* 41 (2018).

[99] W. R. Reader and S. J. Payne. 2007. Allocating time across multiple texts: Sampling and satisficing. *Human–Computer Interaction* 22, 3 (2007), 263–298.

[100] Stuart Russell. 2019. *Human compatible: Artificial intelligence and the problem of control.* Penguin.

[101] Stuart Russell and Peter Norvig. 2002. *Artificial intelligence: A modern approach.*

[102] Stuart J. Russell. 1997. Rationality and intelligence. *Artificial Intelligence* 94, 1–2 (1997), 57–77.

[103] Stuart J. Russell and Devika Subramanian. 1994. Provably bounded-optimal agents. *Journal of Artificial Intelligence Research* 2 (1994), 575–609.

[104] Dario D. Salvucci. 2001. Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human–Computer Studies* 55, 1 (2001), 85–107.

[105] Sayan Sarcar, Jussi P. P. Jokinen, Antti Oulasvirta, Zhenxin Wang, Chaklam Silpasuwanchai, and Xiangshi Ren. 2018. Ability-based optimization of touchscreen interactions. *IEEE Pervasive Computing* 17, 1 (2018), 15–26.

[106] Albrecht Schmidt. 2000. Implicit human computer interaction through context. *Personal Technologies* 4, 2 (2000), 191–199.

[107] Lael J. Schooler. 2004. How forgetting fosters heuristic inference. In *Proceedings of the International Conference on Cognitive Modelling*.

[108] Ben Shneiderman. 2011. Claiming success, charting the future: Micro-HCI and macro-HCI. *Interactions* 18, 5 (2011), 10–11.

[109] Siti Rohkmah Mohd Shukri and Andrew Howes. 2019. Children adapt drawing actions to their own motor variability and to the motivational context of action. *International Journal of Human–Computer Studies* 130 (2019), 152–165.

[110] Satinder Singh, Richard L. Lewis, and Andrew G. Barto. 2009. Where do rewards come from?. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. Cognitive Science Society, 2601–2606.

[111] Mason R. Smith, Richard L. Lewis, Andrew Howes, A. Chu, C. Green, and A. Vera. 2008. More than 8,192 ways to skin a cat: Modeling behavior in multidimensional strategy spaces. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, Bradley C. Love, K. McRae, and Vladimiir M. Sloutsky (Eds.). Cognitive Science Society, Austin, TX, 1441–1446.

[112] Lucy Suchman. 1993. Response to Vera and Simon's situated action: A symbolic interpretation. *Cognitive Science* 17, 1 (1993), 71–75.

[113] Lucy Suchman. 2003. Writing and reading: A response to comments on plans and situated actions. *Journal of the Learning Sciences* 12, 2 (2003), 299–306.

[114] Lucy A. Suchman. 1987. *Plans and situated actions: The problem of human–machine communication.* Cambridge University Press.

[115] Richard S. Sutton and Andrew G. Barto. 1998. *Reinforcement learning: An introduction.* MIT Press, Cambridge, MA.

[116] J. A. Swets, W. P. Tanner, Jr, and T. G. Birdsall. 1961. Decision processes in perception. *Psychological Review* 68, 5 (1961), 301–340.

[117] Kashyap Todi, Jussi Jokinen, Kris Luyten, and Antti Oulasvirta. 2018. Familiarisation: Restructuring layouts with visual learning models. In *23rd International Conference on Intelligent User Interfaces*. 547–558.

[118] Kashyap Todi, Jussi Jokinen, Kris Luyten, and Antti Oulasvirta. 2019. Individualising graphical layouts with predictive visual search models. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10, 1 (2019), 1–24.

[119] Julia Trommershäuser, Laurence T. Maloney, and Michael S. Landy. 2003. Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A* 20, 7 (2003), 1419–1433.

[120] Julia Trommershäuser, Laurence T. Maloney, and Michael S. Landy. 2009. The expected utility of movement. *Brain* (2009), 95–111.

[121] 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.

[122] 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.

[123] Mihaela Vorvoreanu, Lingyi Zhang, Yun-Han Huang, Claudia Hilderbrand, Zoe Steine-Hanson, and Margaret Burnett. 2019. From gender biases to gender-inclusive design: An empirical investigation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[124] Colin Ware. 2019. *Information visualization: Perception for design.* Morgan Kaufmann.

[125] Shumin Zhai, Jing Kong, and Xiangshi Ren. 2004. Speed–accuracy tradeoff in Fitts' law tasks – on the equivalency of actual and nominal pointing precision. *International Journal of Human–Computer Studies* 61, 6 (2004), 823–856.

[126] Jiajie Zhang, Vimla L. Patel, Todd R. Johnson, and Edward H. Shortliffe. 2004. A cognitive taxonomy of medical errors. *Journal of Biomedical Informatics* 37, 3 (2004), 193–204.

[127] Brian D. Ziebart, Nathan Ratliff, Garratt Gallagher, Christoph Mertz, Kevin Peterson, J. Andrew Bagnell, Martial Hebert, Anind K. Dey, and Siddhartha Srinivasa. 2009. Planning-based prediction for pedestrians. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 3931–3936.