

Simulation-based Optimization for Augmented Reading

Yunpeng Bai
National University of Singapore
Singapore

Shengdong Zhao
City University of Hong Kong
Hong Kong, China

Antti Oulasvirta
Aalto University
Helsinki, Finland

Abstract

Augmented reading systems aim to adapt text presentation to improve comprehension and task performance, yet existing approaches rely heavily on heuristics, opaque data-driven models, or repeated human involvement in the design loop. We propose framing augmented reading as a simulation-based optimization problem grounded in resource-rational models of human reading. These models instantiate a simulated reader that allocates limited cognitive resources, such as attention, memory, and time under task demands, enabling systematic evaluation of text user interfaces. We introduce two complementary optimization pipelines: an offline approach that explores design alternatives using simulated readers, and an online approach that personalizes reading interfaces in real time using ongoing interaction data. Together, this perspective enables adaptive, explainable, and scalable augmented reading design without relying solely on human testing.

Keywords

Reading, Modeling, Optimization, User Interface

1 Introduction

Augmented reading aims to adapt text presentation to readers' contexts and capabilities, for example under multitasking conditions [1, 21], time pressure [2], or varying cognitive abilities. Recent systems have explored adaptive summaries, dynamic layouts [20], and AI-assisted transformations of text. However, most existing approaches [17, 25, 26] rely either on heuristic, rule-based adaptations that do not scale across contexts, or on data-driven models [3, 4, 6, 9, 13] that require large amounts of training data and offer limited interpretability. As a result, current augmented reading systems struggle to remain simultaneously flexible, cognitively grounded, and human-centered due to the limitation of existing computational reading models.

We propose that computationally resource-rational models of reading [1, 2] can serve as principled optimization objectives for augmented reading systems. These models conceptualize reading as adaptive decision-making under cognitive and environmental constraints, enabling the simulation of human-like reading behavior across wide situations and reader profiles [19, 23]. Rather than directly predicting interface adaptations from data, which often lack explicit representations of human cognition [3, 4, 6, 9–13], simulated readers can be used to evaluate and compare candidate text presentations from a human-centered perspective, accounting for comprehension, cognitive workload, and time limitations [1, 2]. This reframes augmented reading as a simulation-based optimization problem, where simulated readers represent human readers to act as participants and involve in scalable, adaptive, and explainable text interface design. In this paper, we propose how such

reading models can be integrated into offline and online optimization pipelines, and demonstrate their use through representative augmented reading scenarios.

2 Background: Resource-rational Reading Models

Resource-rational reading models conceptualize human reading as adaptive decision-making under cognitive and environmental constraints [1, 2]. Rather than treating reading as a fixed sequence of perceptual and linguistic stages, these models view reading behavior as emerging from the optimization of task performance given limited resources such as attention, memory, and time. This perspective follows the broader framework of computational rationality [7, 8, 18, 19, 23], which explains human behavior as approximately optimal given internal limitations and external demands.

As illustrated in Figure 1, a reader is modeled as an agent that interacts with a text environment through perception and action. The agent selects reading actions, such as where to fixate, whether to continue or regress, and when to stop based on partial and uncertain observations. Cognitive constraints, including perceptual noise, memory limitations, and attentional capacity, are explicitly represented, inducing uncertainty and trade-offs that shape reading behavior. Reading strategies are derived through optimization in simulation, rather than by directly fitting human data, grounding the resulting behavior in explicit cognitive assumptions.

This simulation-based, resource-rational approach offers several advantages for HCI research [22, 23]. It provides a generative and explainable account of reading behavior, where eye movements, reading speed, regressions, and comprehension emerge from optimization under constraints rather than from heuristics or post hoc data fitting [2]. Because policies are learned in simulation, the approach is also less data-intensive than purely data-driven models while remaining quantitatively testable [27]. Prior work shows that such models reproduce key empirical phenomena in reading and multitasking, including attention switching, resumption costs, and performance degradation under time pressure or divided attention [1, 2, 21]. Crucially for augmented reading, resource-rational models support systematic adaptation across contexts and individuals [2]. Task demands, environmental constraints, and display properties can be modeled through changes in rewards or constraints [1, 2, 14, 21], while individual differences can be captured by adjusting internal parameters such as memory capacity or lexical efficiency [5, 16]. This makes simulated readers suitable not only for explaining behavior, but also for evaluating and optimizing alternative text presentations, enabling principled exploration of adaptive designs beyond static layouts or heuristic rules [2].

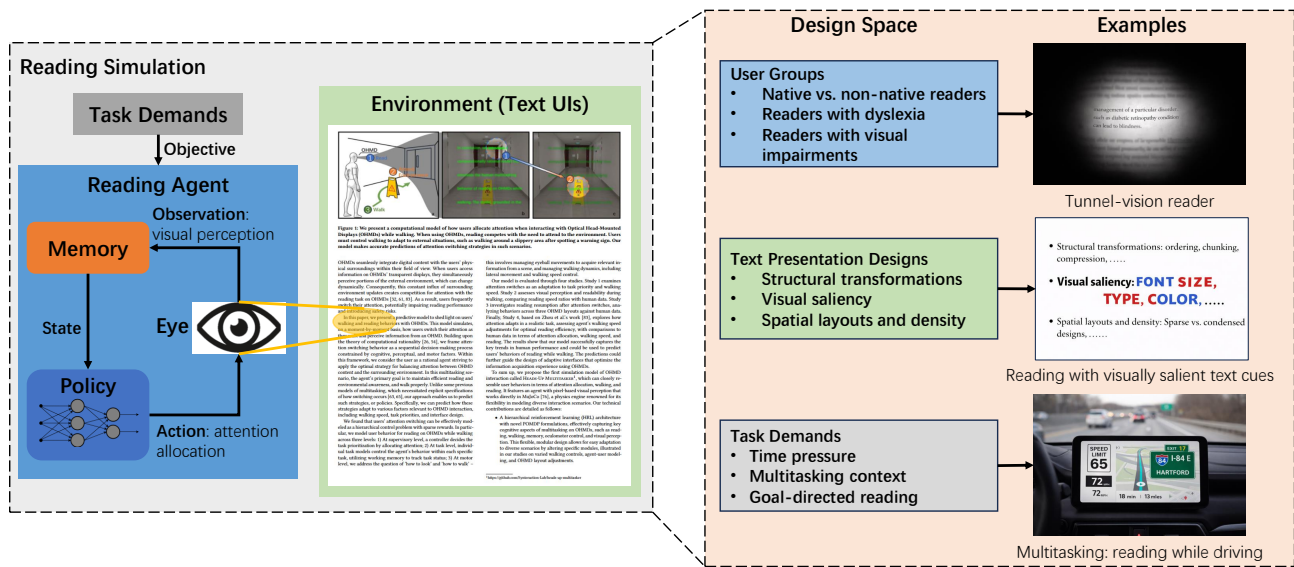


Figure 1: Resource-rational reading models enable principled optimization of text user interfaces across a wide design space. (Left) A reading agent interacts with text-based user interfaces by allocating visual attention over time, guided by perceptual input, memory state, task objectives, and an adaptive policy. (Right) A design space for augmented reading defined by user groups, text presentation designs, and task demands. The same underlying model can be instantiated to evaluate diverse reading scenarios, illustrated here with three examples: (top) a reader with restricted visual field (tunnel vision), (middle) reading supported by visually salient text cues (e.g., font size, type, and color), and (bottom) multitasking contexts such as reading while driving from an in-vehicle display. Together, the model enables principled comparison and optimization of text UI designs across users, tasks, and environments.

3 Augmented Reading as an Optimization Problem

3.1 Definitions

Augmented reading. We define augmented reading as text user interfaces that adapt presentation to support readers under varying tasks, contexts, and capabilities. Rather than treating text as static, augmented reading systems dynamically modify layout, content, or visual properties to improve task performance. Importantly, “good reading” is not defined by a single metric such as speed or readability, but by how effectively readers achieve their goals under cognitive and environmental constraints. Reading involves fundamental trade-offs between comprehension, effort, time, and attention [2, 24]. Augmented reading therefore requires adapting text presentation based on the reader’s situation and task demands, with the goal of supporting task-relevant comprehension as efficiently as possible.

Reading as a resource-rational optimization problem. Resource rationality frame reading as adaptive decision-making under cognitive and environmental constraints. Readers allocate limited resources—such as visual attention, working memory, and time—to maximize task utility, for example comprehension or information extraction [1, 2]. These decisions unfold hierarchically, from eye-movement control to higher-level choices about pacing, rereading, or terminating reading. From this perspective, augmented reading can be formulated as an optimization problem: given a reader, a

task, and a context, how should text be presented to maximize expected utility under resource constraints? This framing shifts attention from individual interface features to the objectives they serve, providing a principled basis for evaluating and optimizing alternative text presentations.

3.2 Design Space in Augmented Reading

Augmented reading systems operate over a rich design space that shapes how text is perceived, processed, and acted upon. As illustrated in Figure 1, this space spans (i) *user groups* (e.g., readers with different perceptual or cognitive capabilities), (ii) *text presentation designs* (e.g., ordering, chunking, compression, visual saliency, and layout density), and (iii) *task demands* (e.g., time pressure, multitasking, or goal-directed reading).

Resource-rational reading models provide a principled way to navigate this space. By explicitly modeling how design factors interact with human cognitive constraints, the same underlying reading model (Figure 1, left) can be instantiated to evaluate diverse scenarios (Figure 1, right), enabling systematic comparison and optimization of text UIs that would be difficult to achieve through manual design alone.

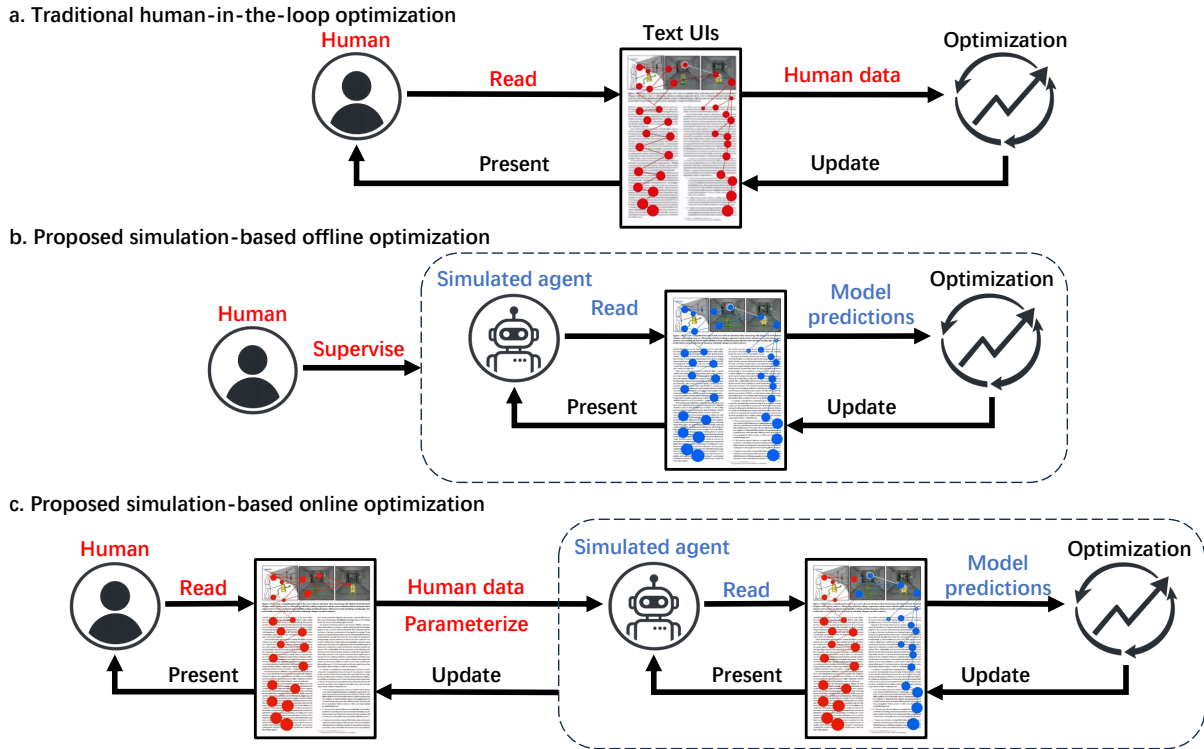


Figure 2: Optimization paradigms for augmented reading. (a) **Human-in-the-loop optimization.** Traditional text UI design relies on human readers to generate behavioral data and feedback, making optimization costly and difficult to scale. (b) **Simulation-based offline optimization.** A simulated reading agent represents human readers to evaluate and optimize candidate text UIs, enabling systematic exploration of the design space without repeated user studies. (c) **Simulation-based online optimization.** Human interaction data are used to initialize a personalized simulated agent that continues reading on the user’s behalf, supporting real-time evaluation and adaptive UI optimization during ongoing reading.

3.3 Resource-Rational Reading Models as Evaluators

Because resource-rational reading models explicitly characterize how human readers adapt to cognitive and environmental constraints, they can be repurposed as evaluators for augmented reading design. Rather than predicting interface adaptations directly from data, as in many data-driven approaches [3, 4, 6], a simulated reader is exposed to candidate text presentations and used to estimate their expected utility in terms of comprehension, effort, and time costs. This evaluator-based perspective enables systematic and consistent comparison of alternative designs across a large design space, supporting exploration and optimization beyond what is feasible through manual iteration or user testing alone [22]. By adjusting internal model parameters such as memory capacity, perceptual noise, or lexical efficiency, the same simulated reader can approximate different user profiles, including readers under extreme multitasking or atypical cognitive constraints. As a result, resource-rational models support reasoning about individual differences and edge cases, enabling more inclusive, robust, and cognitively grounded augmented reading systems.

4 Illustrative Scenarios and Design Implications

We outline two complementary optimization paradigms for augmented reading based on resource-rational reading models: of-line [15, 22] optimization and online [20] optimization (Figure 2). The two paradigms differ in whether humans are inside or outside the optimization loop, but share the same principle: using simulation to guide principled, scalable interface adaptation.

4.1 Offline Optimization

In offline optimization, simulated readers are used to evaluate and refine text interfaces prior to deployment (Figure 2 b). Given a candidate design, such as a layout, summarization level, or saliency scheme, the simulated agent reads under specified task demands (e.g., time pressure or comprehension goals) and produces predicted behavioral and performance metrics, including fixation patterns, reading time, and comprehension. These predictions reveal cognitively demanding regions, such as passages associated with dense fixations or frequent regressions, enabling designers to target revisions (e.g., restructuring content, reducing load, or increasing saliency). Revised designs can be iteratively re-evaluated in simulation, supporting systematic exploration of large design spaces without repeated user studies. By adjusting model parameters, the

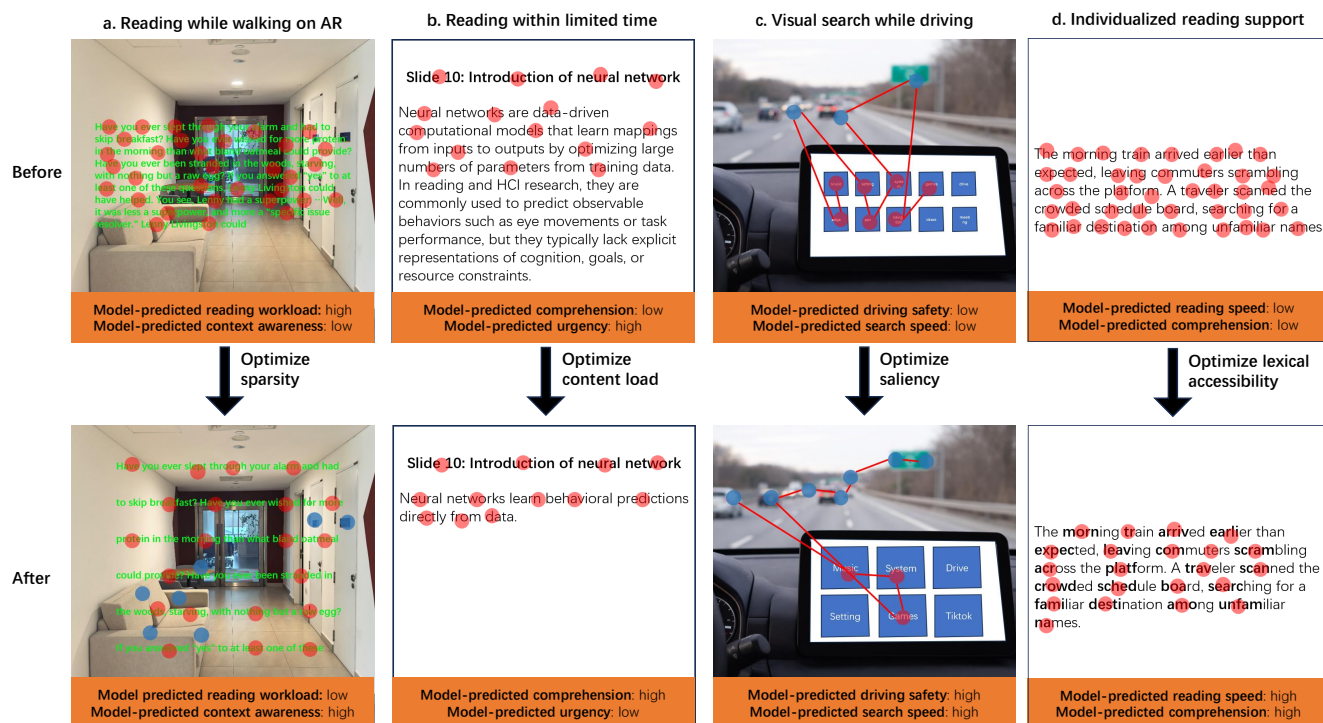


Figure 3: Gallery of simulation-based optimization for augmented reading. Three scenarios illustrate how a simulated reading agent predicts human-like reading behavior (red dots indicate fixations on text; blue dots indicate attention to the surrounding environment) and task performance (summarized by model-predicted metrics in orange boxes) under different contexts, and how these predictions guide UI optimization. Across (a) reading while walking on AR, (b) time-limited reading, and (c) visual search while driving, the model identifies bottlenecks in workload, comprehension, or safety. Targeted design interventions, such as increasing layout sparsity, reducing content load, or enhancing visual saliency lead to improved predicted reading behavior and task performance after optimization. (d) Individualized reading support demonstrates how reader-specific simulation reveals inefficient lexical access (e.g., slow reading speed and dense fixations), enabling personalized adaptations such as enhanced lexical saliency to improve predicted reading speed and comprehension.

same pipeline can approximate different reader populations or capabilities.

4.2 Online Optimization

In online optimization, a simulated reader supports real-time adaptation during ongoing reading (Figure 2 c). Partial observations of a user’s interaction, such as eye movements, reading speed, or task progress are used to initialize or update the internal state of the simulated agent. Conditioned on this interaction history, the agent predicts near-future reading behavior and resource demands under the current text UI. These predictions enable in-situ interface optimization while reading is still in progress. The simulated agent iteratively evaluates candidate adaptations on the user’s behalf, selecting text presentations that better balance comprehension, effort, and time under the current context. By grounding adaptation in an explicit model of human cognition, this paradigm supports dynamic, personalized UI updates that account for high-level cognitive factors rather than relying on surface heuristics or reactive rules.

4.3 Design Implications

Together, these optimization paradigms enable simulation-based design workflows for augmented reading. As illustrated in Figure 3, model predictions can identify task-specific bottlenecks (e.g., workload, comprehension, or safety) and guide targeted design interventions that improve predicted reading and task performance across contexts. This positions resource-rational reading models as practical tools for exploring, comparing, and optimizing adaptive text interfaces.

5 Conclusion

We framed augmented reading as a simulation-based optimization problem grounded in resource-rational models of human reading. By using cognitively grounded simulated readers as evaluators, both offline and online optimization pipelines can systematically improve text interfaces with respect to comprehension, effort, and time under diverse contexts and user profiles. This perspective enables adaptive, explainable, and scalable augmented reading design that

goes beyond heuristics, large-scale data collection, or repeated human testing.

Acknowledgments

A. O. was supported by the European Research Council (ERC; Grant No. 101141916) and the Research Council of Finland (Grant Nos. 328400, 345604, 341763, and 357578). S. Z. was supported by the City University of Hong Kong (Grant No. 9610677).

References

- [1] Yunpeng Bai, Aleks Ikkala, Antti Oulasvirta, Shengdong Zhao, Lucia J Wang, Pengzhi Yang, and Peisen Xu. 2024. Heads-Up Multitasker: Simulating attention switching on optical head-mounted displays. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [2] Yunpeng Bai, Xiaofu Jin, Shengdong Zhao, and Antti Oulasvirta. 2025. A Resource-Rational Mechanism for Reading. (2025).
- [3] Lena S Bolliger, David R Reich, Patrick Haller, Deborah N Jakobi, Paul Prasse, and Lena A Jäger. 2023. ScandL: A diffusion model for generating synthetic scanpaths on texts. *arXiv preprint arXiv:2310.15587* (2023).
- [4] Luca S Bolliger, Daniel R Reich, and Lukas A Jäger. 2025. Scandl 2.0: A generative model of eye movements in reading synthesizing scanpaths and fixation durations. *Proceedings of the ACM on Human-Computer Interaction* 9, 3 (2025), 1–29.
- [5] Suyog Chandramouli, Danqing Shi, Aini Putkonen, Sebastiaan De Peuter, Shanshan Zhang, Jussi Jokinen, Andrew Howes, and Antti Oulasvirta. 2024. A workflow for building computationally rational models of human behavior. *Computational Brain & Behavior* 7, 3 (2024), 399–419.
- [6] Shuwen Deng, David R Reich, Paul Prasse, Patrick Haller, Tobias Scheffer, and Lena A Jäger. 2023. Eyettention: An attention-based dual-sequence model for predicting human scanpaths during reading. *Proceedings of the ACM on Human-Computer Interaction* 7, ETRA (2023), 1–24.
- [7] 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.
- [8] Andrew Howes, Paul A. Warren, Geoffrey D. Farmer, and Wael El-Deredy. 2016. Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review* 123, 5 (2016), 626–647. doi:10.1037/rev0000018
- [9] Yue Jiang, Zixin Guo, Hamed Rezazadegan Tavakoli, Luis A Leiva, and Antti Oulasvirta. 2024. EyeFormer: predicting personalized scanpaths with transformer-guided reinforcement learning. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–15.
- [10] Yue Jiang, Luis A Leiva, Hamed Rezazadegan Tavakoli, Paul RB Housel, Julia Kylvälä, and Antti Oulasvirta. 2023. Ueyes: Understanding visual saliency across user interface types. In *Proceedings of the 2023 CHI conference on human factors in computing systems*. 1–21.
- [11] Yue Jiang, Yuwen Lu, Christof Lutteroth, Toby Jia-Jun Li, Jeffrey Nichols, and Wolfgang Stuerzlinger. 2023. The future of computational approaches for understanding and adapting user interfaces. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–5.
- [12] Yue Jiang, Yuwen Lu, Jeffrey Nichols, Wolfgang Stuerzlinger, Chun Yu, Christof Lutteroth, Yang Li, Ranjitha Kumar, and Toby Jia-Jun Li. 2022. Computational approaches for understanding, generating, and adapting user interfaces. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–6.
- [13] Yue Jiang, Changkong Zhou, Vikas Garg, and Antti Oulasvirta. 2024. Graph4gui: Graph neural networks for representing graphical user interfaces. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [14] Jussi PP 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.
- [15] Francisco Erivaldo Fernandes Junior and Antti Oulasvirta. 2024. AgentForge: A Flexible Low-Code Platform for Reinforcement Learning Agent Design. *arXiv preprint arXiv:2410.19528* (2024).
- [16] Antti Kangasrääsio, Jussi PP 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.
- [17] Gordon E Legge, Timothy S Klitz, and Bosco S Tjan. 1997. Mr. Chips: an ideal-observer model of reading. *Psychological review* 104, 3 (1997), 524.
- [18] 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.
- [19] Falk Lieder and Thomas L Griffiths. 2020. Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and brain sciences* 43 (2020), e1.
- [20] David Lindlbauer, Anna Maria Feit, and Otmar Hilliges. 2019. Context-aware online adaptation of mixed reality interfaces. In *Proceedings of the 32nd annual ACM symposium on user interface software and technology*. 147–160.
- [21] Alexander Lingler, Dinara Talypova, Jussi PP Jokinen, Antti Oulasvirta, and Philipp Wintersberger. 2024. Supporting task switching with reinforcement learning. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [22] Roderick Murray-Smith, Antti Oulasvirta, Andrew Howes, Jörg Müller, Aleks Ikkala, Miroslav Bachinski, Arthur Fleig, Florian Fischer, and Markus Klar. 2022. What simulation can do for HCI research. *Interactions* 29, 6 (2022), 48–53.
- [23] Antti Oulasvirta, Jussi PP Jokinen, and Andrew Howes. 2022. Computational rationality as a theory of interaction. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [24] Keith Rayner. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (1998), 372.
- [25] Erik D Reichle, Keith Rayner, and Alexander Pollatsek. 2003. The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences* 26, 4 (2003), 445–476.
- [26] Dario D Salvucci. 2001. An integrated model of eye movements and visual encoding. *Cognitive Systems Research* 1, 4 (2001), 201–220.
- [27] Richard S Sutton, Andrew G Barto, et al. 1998. *Reinforcement learning: An introduction*. Vol. 1. MIT press Cambridge.