

🔗 Roderick Murray-Smith, University of Glasgow  
Antti Oulasvirta, Aalto University  
Andrew Howes, University of Birmingham  
Jörg Müller, University of Bayreuth  
Aleksi Ikkala, Aalto University  
Miroslav Bachinski, University of Bergen  
Arthur Fleig, University of Bayreuth  
Florian Fischer, University of Bayreuth  
Markus Klar, University of Bayreuth

# What Simulation Can Do for HCI Research

## Insights

- Simulations help create and validate new HCI theory, making design and engineering more predictable, and improving safety and accessibility.
- Emulation of user behavior with generative models tests our understanding of an interactive system.
- Simulation-based intelligence involves directly embedding models in interactive systems.
- Model-based evaluation provides insights into usability *before* user-testing.

The early days of every engineering subject involved examples of expensive failure. The skilled artisans of the day succeeded in making gradual progress, but these successes were punctuated by disasters that occurred when they made too large of an innovation step. From the collapse of cathedrals in France to the capsizing of the 17th-century Swedish warship *Vasa*, to more recent failures such as air accidents attributable to modern cockpit designs, we see the potential high cost of “in the wild” prototyping approaches, especially in modern environments involving rapidly changing demands, or when the complexity and expense of prototyping increase significantly.

While long-established fields such as civil engineering, naval architecture, and aeronautics use modeling and simulation to test designs long before physical prototypes are created, simulation-based design methods are developing rapidly, expanding to new fields such as pharmaceuticals, epidemiology, and medicine, coupling formal models of scientific theory with large-scale data acquisition to calibrate the models. In this article, we argue that simulation can aid in such large technological strides forward but can also support design “in the small,” especially with often-underrepresented user groups or contexts.

Human-computer interaction

research and practice has been slow to adopt simulation, in part because many have argued that traditional human-based usability testing is quicker and more valuable than offline simulation. But the ability to build a generative model that matches user behavior is a strong test of whether we understand an interactive system. Furthermore, simulations can support the creation and validation of new theories, and make design and engineering more predictable and robust processes. Simulation can be directly embedded in intelligent interactive systems, with the potential to improve system safety and accessibility. Model-based evaluation can provide insights into usability *before* testing with end users; we argue that, in the future, in many cases, this cost in money, time, or discomfort of doing extensive parameter optimization via experiments with human participants will make it impractical or unethical to avoid the use of simulation.

## WHAT IS SIMULATION IN HCI?

A model of a system, artifact, or environment is a simplified representation that captures its essential characteristics for a specific purpose. A simulation is the operation of the model, where the intention is to draw conclusions, qualitative or quantitative, about the behavior or properties of a real-world process or system over time. Simulation is an indispensable tool for scientific research that aims to understand the behavior of complex systems, including hypothetical, extreme, or dangerous conditions, or situations where it is too slow or expensive to use the real-world process itself [1]. It allows an unambiguous implementation of the current scientific theory, and predictions can be validated with observed real-world data. Any mismatches prompt researchers to consider the requirements for the next

steps in theory development or data acquisition.

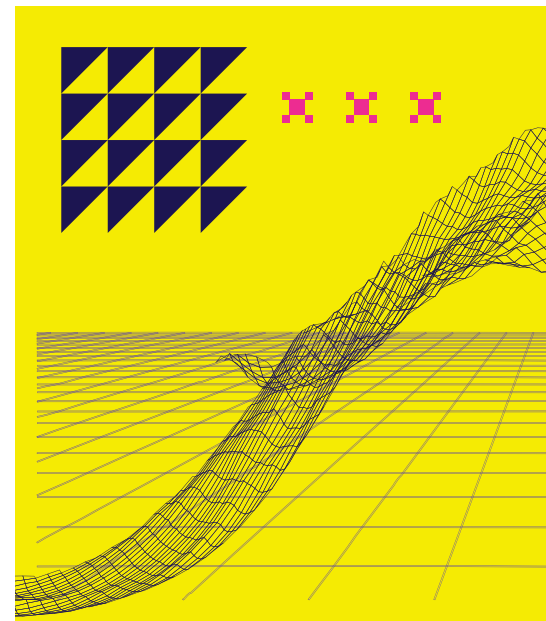
Typically, performing a simulation of a system means using a computer program to approximate the behavior of a mathematical model. In a broader sense, the simulation is a method for: studying systems and their behavior, which includes choosing a model; finding a way of implementing that model in a form that can be executed on a computer; running the model to compute the outputs; using inverse applications of models to infer hidden states or parameters; validating the model; and visualizing, analyzing, and interpreting the resultant data to find explanations.

In HCI, pioneering work in simulations was driven by Card et al. [2], whose model human processor (MHP) divided the user aspect into cognitive, motor-behavioral, and perceptual components. They introduced GOMS (goals, operators, and methods to achieve the goals, selection of competing methods) models to predict task times based on separating tasks into elementary events and summing the expected time to complete the user task.

Criticisms of the use of simulation in HCI have included their cost and complexity to develop, and their inability to adequately represent the cognitive and perceptual complexity of the human in the HCI loop, especially given the sensitivity of behavior to details of context. Other critiques have included the perceived failure of models to capture the physical and social context of interaction. In contrast, we argue that investing in simulation models will actually save expenses and time, by streamlining the development process.

## SIMULATION FOR THEORY FORMATION

Simulation helps push theory forward by virtue of the fact that it demands



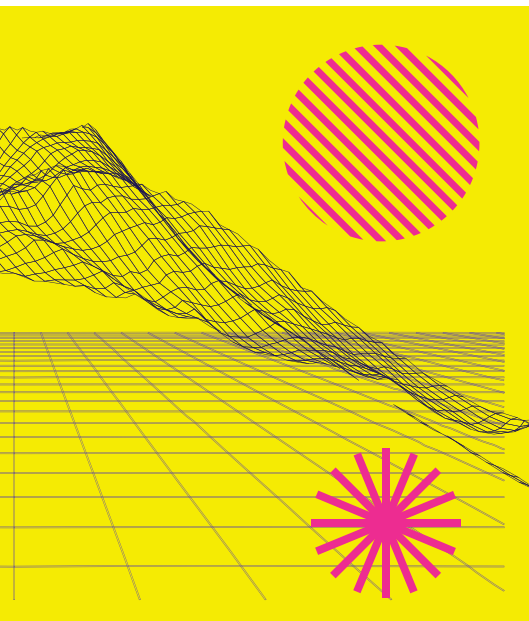
abstraction. Simulation does not require the veridical replication of the movement of every molecule; rather, it must be based on appropriate abstractions that form the “units” of simulation. The choices of these abstractions go hand in hand with theoretical development. In cognitive science, for example, artificial neural networks are often considered to be simple simulations of aspects of biological neural networks. Other phenomena may require different theoretical commitments and radically different kinds of simulators—for example, spiking neural networks.

In HCI and in cognitive architectures, a commitment was made to Fitts’s law because it was felt that stochastic submovements—processes that give rise to the law—were unimportant for predicting pointing performance. Fitts’s law was used as a base-level abstraction in MHP, for example. However, this assumption has proved too restrictive for explaining key phenomena concerning adaptation, and alternative simulation environments are now available. For example, full-body biomechanical simulations using reinforcement learning are used to predict not only movement performance but also motion trajectories and even fatigue during pointing [3].

This coupling of the appropriate abstraction and precision in defining a simulator is also critical in grounding experimental work. As Harold Thimbleby observed:

*We can do as many “experiments” as*

**In the future, this cost in money, time, or discomfort of doing extensive parameter optimization via experiments with human participants will make it impractical or unethical to avoid the use of simulation.**



*we like on complex systems, evaluating systems with vast numbers of people, doing sophisticated statistical tests, and so on, all to no avail unless we know what we are doing, and how the results of the experiment bear on future work.... I search the literature for theories that I can apply in my case...instead I find reports of experiments—sometimes related to my particular problem—but without some underlying theories, how can I know how safely I can generalise those results to apply in my design, with my users? [4]*

However, one lesson that HCI has learned over the past 50 years is that *the details matter*. The context matters, the user matters, and both can change behavior significantly. Understanding and describing the variability of human behavior and sensitivity to context is a significant challenge, but one that can be supported by simulations.

## SIMULATION FOR DESIGN AND ENGINEERING

Can we advance user-centered design to a more rigorous, safer, and predictable process via a simulation-based approach? Simulations can be used to predict task performance, such as time taken to finish a task and how often tasks can be successfully accomplished (for example, the keystroke-level models in Card et al. [2]). They can also be biomechanical models, predicting movement and its physical ergonomics, as well as physiological and health effects. They can have components predicting perceptual performance in different contexts, and can include cognitive

elements. Offline simulations can improve robustness through more-thorough exploration of the design space.

Many interactions have already been modeled with biomechanical models, from a button press to the touch on a screen to midair gestures to full-body movements [3]. These interactions can be evaluated by comparisons with human data, for example, from motion capture. Simulations of such movements are computationally demanding, but today various suites of dedicated biomechanical simulation software can be found to efficiently carry out these computations, including OpenSim (<https://simtk.org/projects/opensim>), AnyBody, LifeModeler, and SantosHuman, as well as powerful physics engines with biomechanical modeling capabilities such as MuJoCo (<https://mujoco.org>) and Bullet (<http://bulletphysics.org>). Biomechanical simulations can include *inverse simulations*, which enable the inference of hidden states or parameters from experimentally observed movement data, and *forward simulations*, which predict complete movement behavior.

It will often be impossible to avoid simulations because of the cost both financially and timewise of doing extensive parameter optimization via experiments with human participants. In Williamson et al. [5], for example, a bearing-based pedestrian navigation system had parameters such as the size of the angular window needed for feedback when pointing at the target. A simulator modeling pedestrians as rational agents optimized the ideal window size for efficient navigation under different assumptions of sensor uncertainty—an extremely time-consuming task for multiple large groups of human participants. Per Ola Kristensson and Thomas Müllners [6] use modeling to replace extensive experimentation, thus optimizing text-entry system parameters.

## SIMULATION FOR INTELLIGENT INTERACTIVE SYSTEMS

Previous sections described the role a simulation can have *offline*, during the design stage. However, simulations can also run *online*, in real time, or faster than real time, with their initial states based on current conditions. This can

allow a system to act as a “digital twin,” monitoring activity and inferring hidden states (such as possible user intentions or goals), or it can predict possible outcomes and use this information to adapt the interface, make decisions, or change the feedback to the user.

The ability to perform faster-than-real-time simulation allows predictive interfaces to offer auto-complete options, or to jump to likely targets, or permit more-sloppy actions from the user [7]. In more safety-critical applications, it could be used to provide the user with warnings about the consequences of their current behavior, if there is a possibility that it will lead to dangerous states.

## SIMULATION FOR SAFETY, ACCESSIBILITY, AND ETHICAL REASONS

Traditional user testing has its place, but as systems become more complicated, and users more diverse, we hit complexity, robustness, and ethical challenges in usability testing. Melissa Quek observes, “In contrast, it can be more difficult to access people with disabilities and a user study can take longer and is more effortful for the participant. Inclusive design aims to make it possible for mainstream applications to be used by people of all abilities. To this end, good models and guidelines must be made available to designers and developers” [8]. In the near future, it may be deemed unethical in some domains, such as those with vulnerable users, to propose an experiment with human users before every effort has been made to reduce the uncertainty about the outcome with other means. A key element will be a rigorous simulation of the experiment.

Simulation can be necessary for a number of reasons. It may be too risky to test a system without initial simulation of proposed experimental conditions. The risks can be physical, emotional, or ethical. For niche user groups, the availability of users locally may be very limited, and it may be difficult to persuade participants to take part in multiple trials (or ones not in line with experimental protocol). Use of simulation forces designers to be explicit about the elements included in the model, making the design process more auditable for stakeholders; for

example, to identify underrepresented users.

When designing for inclusion, an empathic modeling approach focuses on simulating a disability to allow designers to understand a system from the user's point of view and to better appreciate the problems the system should tackle, allowing a narrowing down of options and limiting fatigue and frustration for the participants. Quek [8] provides a review of the literature of simulation as part of the evaluation process for vulnerable groups. She also provides a specific example of using simulations of brain-computer interfaces to explore design options and allow able-bodied users to test the interaction before disabled users were asked to test the system. This approach also allowed the creation of multiparticipant software, such as games where people of different input abilities and input mechanisms could be placed on an equal footing by using the simulations to create a common denominator among all users.

For safety reasons, simulation has long played a key role in aviation in training pilots and testing new flight procedures or aircraft design changes. Similarly, as autonomous driving and associated interfaces grow in importance, we anticipate an expanded need for simulation in UI design for automobiles. In general, deploying untested systems to millions of users is highly risky in terms of reputation, customer retention, and longer-term financial consequences. In recommender systems, for example, a gap has opened between research and practice, due to the vulnerability of the traditional approach of testing on historical logs of user interactions. Recommender systems that perform well on historical data often rapidly go wrong when they engage with real users. While A/B testing with population subsamples can reduce risks, closed-loop simulations based on

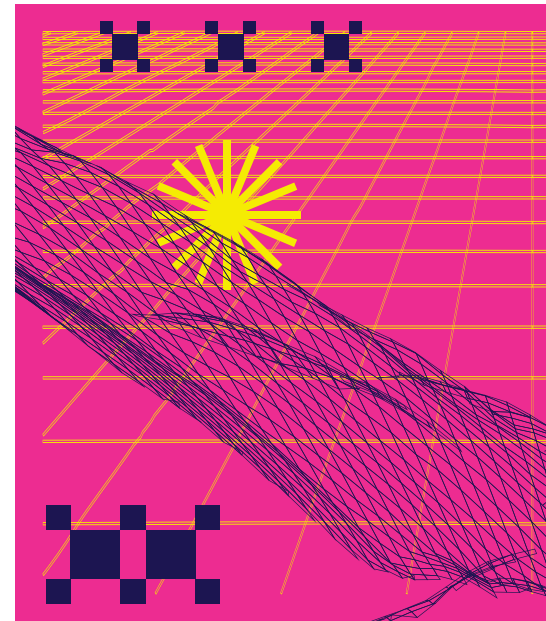
user models can be used to pretest the system before user testing.

## WHY NOW?

Why do we believe that the time is ripe to reconsider what models can do for HCI? Lavin et al. [1] present a far-reaching topical review of the role of simulation in science and AI (but few HCI examples), highlighting that in the past, the complexity of simulation was constrained by hardware limitations, lack of information, the difficulty of dealing with uncertainty, and practical challenges in integrating multiple different simulation models, reducing the utility of the simulation approach for practical decision making. However, recent developments in probabilistic, differentiable programming, high-performance computing, and causal modeling and the rapidly improving ability of machine learning to emulate complex aspects of human perception and behavior mean that simulation has an increased potential to be efficiently and usefully applied in new domains such as HCI [9].

New sensor technologies and new interaction styles, such as augmented reality, take us out of our comfort zone with well-understood mechanisms, due to issues such as sensor fusion, high-dimensional and uncertain sensors, and the application of machine-learning technology for the segmentation and labeling of content. The increased complexity of system design will demand more use of simulation in its development and optimization.

A common challenge, relevant to HCI, is linking simulation-based models with empirical data. Numerical simulators typically have parameters whose values are not known a priori and have to be inferred by data. Classical statistical approaches are not always easy to apply to models defined by numerical simulators, and in some cases simulations may be in legacy



code, or only available as black boxes. A key recent development is *simulation-based inference* (SBI), also known as *likelihood-free inference*, which enables researchers to algorithmically identify parameters of simulation-based models that are compatible with observed data and prior assumptions.

## OUTLOOK

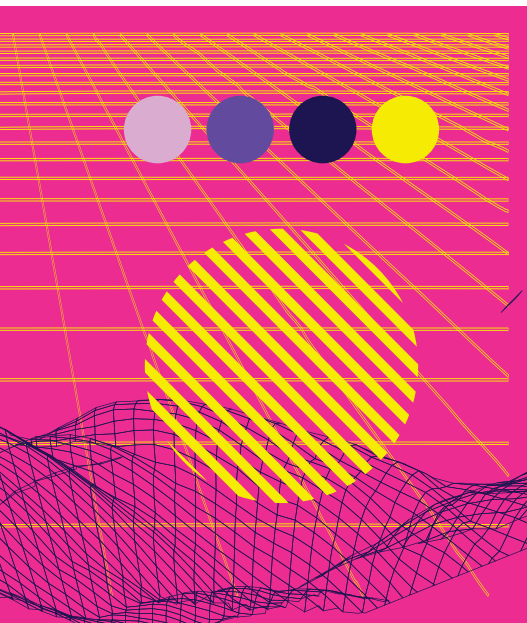
We have highlighted that developments in hardware and software can enable us to create increasingly complex models of human interactions with technology, and calibrate them to increasingly rich and available data. An example is the application of methods from artificial intelligence and machine learning, which are having a wide-reaching impact on many areas of day-to-day life and in science. We believe that the natural outcome of the application of these technologies will be in simulation systems that can have a combination of first principles, white-box modeling, and flexible, data-driven black-box models. It is important that stakeholders with an understanding of specific user groups be involved in the specification and evaluation of these models.

Simulations can be used offline, during the design process, to reduce stress on vulnerable user groups, increase the rigor and reproducibility of testing, ensure diversity in testing processes, and improve safety. They can also speed up design and development and reduce project development time uncertainty.

Use of simulations online allows us

---

**When designing for inclusion, an empathic modeling approach focuses on simulating a disability to allow designers to understand a system from the user's point of view.**



to include a predictive element in the computational design of interfaces, and explore multiple scenarios compatible with observed data. Such online simulation of human behavior is likely to be a core requirement of any future “intelligent” interactive systems.

In addition to these practice-focused improvements, simulation can help the scientific process in HCI research. The need for formal rigor in the creation of a simulation model, and for controlling and documenting the provenance of data used to calibrate it, makes clear the importance of many of the often poorly described aspects of context in HCI experiments. A simulation package is also easily shared with other researchers, improving reproducibility.

Aspects of models that at any given moment in time are poorly justified theoretically, are a poor fit to experimental data, or are highly sensitive to context can be viewed as prompts to the research community about where they need better theories, more complex models, or more data. This can create a shared awareness of the open problems and challenges, and can help document progress.

Lord Kelvin said, “When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind.” We argue that if we cannot build a generative simulation model that can,

to some useful degree, replicate aspects of the complexity and variability of human behavior in a given interaction context, then our knowledge of the expected interaction and its consequences is still of a “meager and unsatisfactory” kind. If we have a concrete simulation model with known weaknesses that need improving, however, then at least we know where to begin to develop our theory and acquire more data, to rectify that unsatisfactory state of affairs.

#### ENDNOTES

1. Lavin, A. et al. Simulation intelligence: Towards a new generation of scientific methods. *arXiv preprint arXiv:2112.03235*, 2021.
2. Card, S.K., Moran, T.P., and Newell, A. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum, 1983.
3. Cheema, N., Frey-Law, L.A., Naderi, K., Lehtinen, J., Slusallek, P., and Hämäläinen, P. Predicting mid-air interaction movements and fatigue using deep reinforcement learning. *Proc. of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2020, 1–13.
4. Thimbleby, H. *User Interface Design*. ACM Press, New York, 1990.
5. Williamson, J., Robinson, S., Stewart, C., Murray-Smith, R., Jones, M., and Brewster, S. Social gravity: A virtual elastic tether for casual, privacy preserving pedestrian rendezvous. *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 2010. DOI: 10.1145/1753326.1753548
6. Kristensson, P.O. and Müllners, T. Design and analysis of intelligent text entry systems with function structure models and envelope analysis. *Proc. of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2021, 1–12.
7. Weir, D., Pohl, H., Rogers, S., Vertanen, K., and Kristensson, P.O. Uncertain text entry on mobile devices. *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 2014, 2307–2316.
8. Quek, M. *The Role of Simulation in Developing and Designing Applications for 2-Class Motor Imagery Brain-Computer Interfaces*. Ph.D. thesis, University of Glasgow, 2013.
9. Jokinen, J.P.P., Sarcas, S., Oulasvirta, A., Silpasuwanchai, C., Wang, Z., and Ren, X. Modelling learning of new keyboard layouts. *Proc. of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2017, 4203–4215.

➦ **Roderick Murray-Smith** is a professor in the School of Computing Science at the

University of Glasgow, where he is a member of the Inference, Dynamics and Interaction group and works in the areas of human-computer interaction, machine learning, and control.  
→ [Roderick.Murray-Smith@glasgow.ac.uk](mailto:Roderick.Murray-Smith@glasgow.ac.uk)

➦ **Antti Oulasvirta** leads the User Interfaces research group at Aalto University and the interactive AI research program at the Finnish Center for AI.  
→ [Antti.Oulasvirta@aalto.fi](mailto:Antti.Oulasvirta@aalto.fi)

➦ **Andrew Howes** is a computer science professor at the University of Birmingham interested in the application of computational thinking to explain human behavior. He is also interested in how to design tools that help people make better decisions.  
→ [A.Howes@bham.ac.uk](mailto:A.Howes@bham.ac.uk)

➦ **Jörg Müller** is a full professor of computer science in human-computer interaction at the University of Bayreuth, Germany. His research interests include modeling, simulation, and optimization of HCI using dynamical systems models, biomechanical simulation of HCI, ultrasonic levitation interfaces, and augmented and virtual reality.  
→ [Joerg.Mueller@uni-bayreuth.de](mailto:Joerg.Mueller@uni-bayreuth.de)

➦ **Aleksi Ikkala** is a computer scientist with a background in cognitive neuroscience. He is curious about the possibility of creating intelligent machines, either by developing the cognitive abilities of machines as such or by augmenting biological objects with, for example, neuromorphic parts.  
→ [aleksi.ikkala@aalto.fi](mailto:aleksi.ikkala@aalto.fi)

➦ **Miroslav Bachinski** is an associate professor at the University of Bergen. His research focuses on the development and application of data-driven methods to improve post-desktop user interfaces within the large spaces of alternative designs (e.g., virtual reality, levitation).  
→ [Miroslav.Bachinski@uib.no](mailto:Miroslav.Bachinski@uib.no)

➦ **Arthur Fleig** is a postdoctoral researcher at the University of Bayreuth. His research areas include optimal control methods for human-computer interaction—especially using model predictive control and deep reinforcement learning—acoustic levitation, and the Fokker-Planck Equation.  
→ [arthur.fleig@uni-bayreuth.de](mailto:arthur.fleig@uni-bayreuth.de)

➦ **Florian Fischer** is a research associate at the Chair of Serious Games at the University of Bayreuth, with a background in mathematics. His research areas include optimal control methods for human-computer interaction, deep reinforcement learning, system identification, and IRL.  
→ [florian.j.fischer@uni-bayreuth.de](mailto:florian.j.fischer@uni-bayreuth.de)

➦ **Markus Klar** is a research associate at the Chair of Serious Games at the University of Bayreuth. His research areas include optimization of dynamic interaction techniques, model-predictive control, and simulation of human-computer interaction.  
→ [markus.klar@uni-bayreuth.de](mailto:markus.klar@uni-bayreuth.de)