Does HCI Scale? Scale Hacking and the Relevance of HCI

Barry Brown, Stockholm University

Susanne Bødker, Aarhus University

Kristina Höök, Royal Institute of Technology (KTH)

Insights

- The use of design fictions, mid-range theories, software and hardware populations, and data-driven design tools and research methods is a start on addressing scale.
- Is addressing scale possible without access to millions of users or millions of different systems?

HCI has had a massive impact on the world through streamlining and enabling millions of interfaces on billions of devices. As we face the potential of a tenfold increase in the number of devices and their complexity, it is worth asking about the relationship between HCI and scale. Do the tools and research methods we currently deploy scale to the millions of future interfaces and systems, used by billions of people, across multiple contexts? In this article we outline how we see the challenge of scale. By *scale* we mean how technology is used in large networks of interconnected systems, with billions

of users, across diverse contexts. How can we understand and design for this complex of interconnected uses? Put simply, does HCI scale?

We discuss three different scales: the number of users, the different contexts of use, and the multitude of systems and technologies. Emergent phenomena like filter bubbles or “fake news” can exist only when millions of people are sharing and interacting online on a daily basis. While approaches like statistical generalization offer the possibility to predict the reliability of certain phenomena, they can do that only with phenomena that already exhibit themselves—predicting and

understanding new phenomena is much harder. In turn, while we can examine and generalize from current-day, large-scale system use (such as studying Wikipedia or Facebook), this does not tell us how these systems change over time. In contrast, qualitative analysis can detect future trends or help us to understand and design new systems, but this too can be hampered by the difficulty in predicting what is fundamental (and will remain) and what is passing or unimportant.

Focusing on HCI, how do prevailing popular HCI research methods, in particular research through design (RtD), critical design, interface-technique-driven development, and data-driven design methods, address work at scale? We suggest here four different “hacks” that can help with the challenges to these methods.

WHAT IS SCALE?

There are many ways in which HCI has moved beyond one user/one device, and others where it has not yet done so. One part of this puzzle is being able to reason *across different settings*. More and more, technologies are being applied across contexts, such as when work technologies are brought into the home, or when people use technologies in different places or in different forms of collaboration with others. Mobile devices are perhaps the most obvious examples of this, but social media, ubiquitous text messaging, eHealth, and even casual gaming have insinuated software into its current almost ever-present form. While of course there are studies of each of these, the challenge of scale here is that we still lack an understanding across these different contexts. Indeed,

when we see hybrid forms (such as, say, Snapchat’s combination of photography and messaging), HCI is usually rather late to understand them and thus needs to develop its understanding afresh. What then are the tools that can help us understand systems across different settings? Universal technologies may be one answer, yet other research points to specific use activities where the context-crossing has often been ignored.

A related issue comes when we need to reason about usage where the *number of users* expands beyond what can be considered for a practical trial or study. In experimental and explorative HCI, there is a tendency to focus on a limited number of users, but this offers little help when scaling to mass usage. While statistical generalizations can give confidence in phenomena that manifest themselves at this small scale, they cannot identify phenomena that would occur only at the bigger scale.

In turn, there is a tendency in interaction design to focus on *one technology at a time*. We design the piece but apply it in a whole, into contexts where there are already other artifacts and systems. For research this means that we need to address the wholes of infrastructures and artifact ecologies and not simply the singular systems and designs. Indeed, the innovative systems we study today are perhaps like jigsaw-puzzle pieces, the significance of which will not be apparent until other parts of the puzzle are adopted. How then can we understand a single piece, without access to the whole puzzle? The application in the whole may not simply be a smooth transition of the artifact ecology. As a matter of fact, a new design may disrupt the ecology in ways that cause tensions and more profound changes, perhaps on a longer time scale (see, e.g., [1]).

Indeed, ecologies of artifacts, services, and data may lead to unexpected and unpredictable interactions. *Layers of systems* on top of one another may also relate to scale: Through a touch of the iPhone, we may change the heating or lighting of a whole house, steer a ship, or move an atom in a quantum computer. In this manner, that touch—the human-computer interaction—cuts across human scale as well as super- or supra-scales [2].

Layers of systems on top of one another may also relate to scale: Through a touch of the iPhone, we may change the heating or lighting of a whole house, steer a ship, or move an atom in a quantum computer.

SCALE IN HCI

Given these challenges to scale—in size of data and in ecologies of artifacts—

we need to ask how well HCI research methods address them. Do our methods *explicitly* handle big data? Sometimes this can be addressed by moving back and forth between quantitative and qualitative data in rapid succession, designing and redesigning services that rely on big data (such as Spotify or Facebook). For other problems, we might address scale *implicitly* if the method guarantees that questions are asked that allow it to scale beyond the specific, limited setting that is empirically examined.

In turn, it is not enough that research methods can handle many users or go across many devices. The *outcome* of the research process—the knowledge contribution—also has to scale and be applicable to many different design situations, contexts, domains, and so on. We characterize these attempts as *scale hacking*—ways of conducting research that can go beyond the particular situations being addressed or studied and be generalized in new ways.

Scale has of course always been a concern to HCI researchers. There has, for example, always been work that relies on statistics and quantitative data to statistically generalize over populations. In the early days of HCI, through its origins in ergonomics and cognitivism, HCI researchers attempted to find behavior patterns that would be stable over different contexts, independent of specific technological solutions. Fitts's law is one example; creating software and/or hardware that is made available as open source tools is another. With software, this is relatively easy, as one line of code can be duplicated and run on thousands or even millions of devices, and in turn be run millions of times.

FOUR SCALE HACKS

We use the term *scale hacks* here, as our suggestions are not complete methods in their own right but rather are opportunistic modifications of existing methods to deal with scale. As one would expect with hacks, these are not complete solutions. Instead, they are more a way of beginning to deal with and engage with the problems that scale presents. First, can we think about the role that culture plays in design? Second, can we think across interaction techniques to broader interaction *gestalts*? Third, how can we change design to make use of data? Last, what

role can longitudinal studies play in letting us understand scale?

Going beyond individual designs.

Research through design (RtD) and various forms of critical design are popular methods in contemporary HCI. We have seen a lively debate on ways of framing design knowledge that can transcend different technological waves or design contexts (sometimes framed as mid-range theories, such as strong concepts or experiential qualities [3]). We have also seen a lively debate on how to frame particular critical designs against the backdrop of a longer history—the so-called critical thread of research. By putting particular critical designs into a context of, for example, art styles or whole traditions of work in the humanities, the effect of a particular critical design project, and its contribution to a high-level, ongoing debate, might scale [4].

Sometimes the ideas arising in a design-driven project are organized in whole methodologies, providing particular design judgments or value-based approaches to design work. An example is the well-known Utopia project, originally aimed at shaping desktop publishing with laser printing in a direction that would keep the sensitive, skill-based aesthetics that typesetters had, but that instead brought us the whole idea of participatory design with its strong political values [5]. What is perhaps interesting here is that we seldom discuss how ideas are generated in HCI, and how these *ideas* might have purchase. We find discussions that address “motor themes” or concepts like interaction somehow unsatisfactory, in that they are attempts to rather bluntly explore the sorts of topics we address, rather than thinking about ideas more centrally.

When we discussed our definition of scale above, we realized that HCI suffers from relatively stunted ideas of how culture at scale can influence the development of technology, and how technology at times develops in quite unexpected directions. For something like touch technology, for example, the extent to which it was of interest to HCI was as an input technology for large displays. In reality, clearly its true home was in enabling input to smaller displays. How could we have foreseen this misdirection? How can we think in innovative ways about the connections

between particular HCI work and the world in which it must ultimately find a home? One promising, but clearly partial, technique is the use of design fiction. Design fiction is an approach in which fictional accounts of the use or presentation of particular technologies are used to envision a new future.

In our own work, for example, we created a “future IKEA catalog” in which unusual and inventive different types of IKEA devices were advertised and listed in an imaginary IKEA catalog from the future [6]. Our goal in doing this was to think a little more about how IKEA furniture could be hybridized with innovative new technologies, but also to think about how the world, business models, and methods that IKEA works with would change and influence that technology in development. To take one example, IKEA as a company has always reduced costs by moving work onto the user, such as transporting furniture from the store to home, assembling furniture, and in some cases even fitting existing parts of furniture together to create a user's own arrangement. If there are 3D printers or devices that in some way can produce raw materials, we might think about how IKEA might simply offer users plans that they themselves can print, rather than relying upon the IKEA production process. Here we take a new technology and consider how it will develop in IKEA's world, with the kinds of constraints on its production and development with which the company is familiar.

Of course, one rather serious problem with design fiction is that, as a fiction, it is likely, perhaps fatally, to replicate the prejudices and issues of its authors, and it perhaps lacks the generality that comes from data that might contradict those wishes in some way. Moreover, like fiction, much of it is at times poor. But we see at least the *possibility* of being able to scale in new ways with design fiction, to think about longer-term processes that take design innovations but twist them into new shapes. One way to challenge the design-fiction method could be to take a scale perspective. That is, using fictions to ask questions like: Would this scale to millions of users and still be sustainable and ethical? If there were to be many interfaces and interactions like these, what social or societal changes could follow?

Ecologies of software and artifacts. User interface software and technology (UIST) is concerned with a broad range of technical issues in interfaces, such as tools for user interface development, advanced interaction technologies thriving off of new materials, and software architectures of interactive systems. While UIST research engages with the problems of understanding the user's position when systems scale, scale presents challenges there with the complexity of systems when they scale, and the role of humans in this.

One interesting approach is work on software populations, and studying software ecologically as we might study animals in an environment. Studying software as a population focuses on how different software interacts in specific environments, and even

replicates and spreads, sometimes in autonomous ways, or at least autonomous to the software developers [7]. Large-scale system design has to some extent addressed these issues, but with much less attention to the case where the software is produced by different authors, and where systems have complex interactions between the different parts. Thinking about software (and users) in this way raises a number of interesting questions. Not much HCI work has focused on the interaction between apps, even in the hands of one user. In the contemporary situation where there are multiple apps with similar functions available in app stores, we encourage studies of how such apps interact, how one gets chosen over another by users, and how users arrange their uses of several apps over time and activity, not to mention how

app stores as such mediate the choice and sharing of apps [8].

If we think of the longer-term deployment of applications, we might also reflect on their life cycles—in particular the ways in which they may be updated to new versions, or how applications might die or be reborn in new forms. To add to this complexity, there are also questions of how applications might depend on each other or create ecological niches for other applications. The markets in filters for Photoshop, for example, or new software keyboards for mobile phones, exist as niches in the world that is created by other software products.

Data as a design material. When discussing how to make HCI research methods scale, data-driven research methods are perhaps the first to come to mind. Data is often discussed in terms of its potentially enormous value in modeling and predicting behavior. This in turn can be used to create various actuations in the world: controlling processes in industry, organizing logistics, handling so-called smart cities and smart homes, and augmenting our bodies. It seems clear to both academia and industry alike that patterns in data will be the basis of a whole thriving field of innovations, novel services, and novel interactions. But in order to serve that role, it is not enough to just engage with techniques of statistical generalization or inference. Rather, we are thinking about how data can support the actual intellectual process of idea generation. While this might seem like a suspect idea, in the world of design there is increasing use of machine-learning techniques to generate possible design solutions.

Simply gathering data without knowing what actuation that data can drive is, if not meaningless, at least misguided. The design process that makes use of data as a material is not one of finding some unknown pattern in a preexisting dataset that surprises the design team, allowing them to innovate new functionality. The choice of data sources is not objective, but rather constitutes a choice in the design process that in turn will determine what services or technical actuation can be done. It is a chicken-and-egg problem.

Once there is a first application where data is used in interesting ways,

It is not enough to just engage with techniques of statistical generalization or inference. Rather, we are thinking about how data can support the actual intellectual process of idea generation.

such as in the early experiments with movielens.org or Minecraft, then new functionality and new possible interactions can be imagined. For example, the crowdsourced city planning in Minecraft could not have been imagined before Minecraft existed as a multi-player game. In our view, these exemplify interesting processes of crowdsourcing design in and through massive amounts of data. Partially, though, this gives us some arguments for attempts (such as app-store trials) to expand the user base involved in particular trials, thus overcoming the bootstrapping problems. It may be that certain phenomena appear only when systems have a certain complexity or a certain size, and the right combination will unlock interesting results. HCI researchers have occasionally made sure their designs can find a significant audience through the development of proper mobile apps or Internet-based applications [8]. Once you have a stable group of users, you can test theories by altering the design for one group of users and then statistically comparing the behavior changes.

Scaling over longer time periods.

Last, we might think about how we can simply continue or expand our existing methods but prolong them through longer studies or more industrial-style processes. This might be characterized as “trying harder”—simply taking what we already do but attempting to prolong or expand its scale. One example of this is the exploration of batch design methods, where they produce a hundred versions of one artifact rather than just a few [9]. By expanding the length or scale of a particular technology trial, there is the possibility that we might be able to gain a more diverse perspective on use.

One challenging question here, though, is whether this does actually address the analytic problem of scale. What we are looking for is not just more of the same but instead to actually get a handle on the sorts of issues that are generated when there are larger interactions (be they between users or different systems). The challenge then is to pass on doing existing trails with more people in favor of systems where the larger number of users has a material effect on what each user does. One interesting example of this

is Kraut et al.’s Homenet studies run in Pittsburgh, which pioneered the deployment of Internet technology [10]. Two innovative aspects of this work are that the original deployments were of sufficient scale for complete communities to be established, hence generating new types of interactions between participants. The Homenet trial ran for 20-plus years—to the point where the Homenet systems were superseded by market-based Internet technology (at least partially). This longitudinal component gives us one of the few cases where we can contrast how market innovations can replicate but also fail when compared with academic-produced innovations.

CONCLUSION

Our original question was: What are the tools/research methods we can use to reason about millions of interfaces, thousands of systems, used by billions of people? Clearly our suggestions here do not fully address this question. Adding design fictions, mid-range theories, software and hardware populations, or data-driven design tools and research methods is only a flawed start. Perhaps if we had access to millions of users and millions of interfaces under our control, we could ask really interesting research questions and then use quite different research methods to figure out the answers. Yet for most research studies, this is beyond what is possible or practical. This means that we need our existing methods, and new methods, to scale up. We should not be content simply to figure out limited domains or sites. Instead, we should ask more generic questions about how best to integrate new things into the ecologies of artifacts we already live with. What is needed are new research methods that “talk back,” provide alternative futures, and push back on our thinking, instead of scaling by just adding access to data or ecologies of artifacts.

ENDNOTES

1. Bødker, S. and Klokmoose, C.N. Dynamics in artifact ecologies. *Proc. of NordiCHI'12*. 448–457.
2. Bødker, S. and Andersen, P.B. Complex mediation. *Human Computer Interaction* 20, 2 (2005), 353–402.
3. Höök, K. and Löwgren, J. Strong concepts: Intermediate-level knowledge

in interaction design research. *ACM Transactions on Computer-Human Interaction* 19, 3 (2012), 23.

4. Bardzell, J. and Bardzell, S. Humanistic HCI. *Synthesis Lectures on Human-Centered Informatics* 8, 4 (2015), 1–185.
5. Kyng, M. Bridging the gap between politics and techniques: On the next practices of participatory design. *Scandinavian Journal of Information Systems* 22, 1 (2010), 49–68.
6. Brown, B., Bleecker, J., D’Adamo, M., Ferreira, P., Formo, J., Glöss, M., Holm, M., Höök, K., Johnson, E.C.B., Kaburuan, E., and Karlsson, A. The IKEA Catalogue: Design fiction in academic and industrial collaborations. *Proc. of the 19th International Conference on Supporting Group Work*. ACM, 2016, 335–344.
7. Higgs, M., Morrison, A., Girolami, M., and Chalmers, M. Analysing user behaviour through dynamic population models. *CHI 2013 Extended Abstracts*. ACM, 271–276.
8. McMillan, D., Morrison, A., Brown, O., Hall, M., and Chalmers, M. Further into the wild: Running worldwide trials of mobile systems. *Proc. of the International Conference on Pervasive Computing*. Springer Berlin Heidelberg, 2010, 210–227.
9. Gaver, W., Boucher, A., Jarvis, N., Cameron, D., Hauenstein, M., Pennington, S., Bowers, J., Pike, J., Beitra, R., and Ovalle, L. The Datacatcher: Batch deployment and documentation of 130 location-aware, mobile devices that put sociopolitically-relevant big data in people’s hands: Polyphonic interpretation at scale. *Proc. of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, 1597–1607.
10. Kraut, R., Scherlis, W., Mukhopadhyay, T., Manning, J., and Kiesler, S. HomeNet: A field trial of residential Internet services. *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1996, 284–291.

📌 **Barry Brown’s** recent research has focused on the sociology and design of leisure technologies—computer systems for leisure and pleasure. In recent years he has used video methods to study mobile phone and Apple Watch use; he is currently working on an “at scale” study using similar methods.
→ barry@dsv.su.se

📌 **Susanne Bødker** is well known for her early work on participatory design and then later on for activity theory, as well as for framing the third wave of HCI.
→ bodker@cs.au.dk

📌 **Kristina Höök** is active in interaction design and is currently focusing on somaesthetic design and issues around Internet of Things development. Her earlier work focused on social navigation, seamfulness, and affective interaction.
→ khook@kth.se