

Introduction to HCI research

1

1.1 INTRODUCTION

Research in the area of human-computer interaction (HCI) is fascinating and complex. It is fascinating because there are so many interesting questions and so many changes over time (due to technical advancements). It is complex because we borrow research methods from a number of different fields, modify them, and create our own “standards” for what is considered acceptable research. It is also complex because our research involves human beings who are, to put it mildly, complex. It is important to understand the roots of the field, to understand the development of research methods in HCI, understand how HCI research has changed over time, and understand the multiple dimensions that must be considered when doing HCI research.

1.1.1 HISTORY OF HCI

There is a general consensus that the field of HCI was formally founded in 1982. This is the date of the first conference on Human Factors in Computing Systems in Gaithersburg (Maryland, United States), that later turned into the annual ACM SIGCHI conference. So, at the publication time of this book (2017), the field of human-computer interaction (HCI) is around 35 years old. However, this is a deceptively simple description of the history of HCI. The field draws on expertise existing in many other areas of study. People were doing work before 1982 that could be considered HCI work. There is a fascinating article (Pew, 2007) that describes work on a project for the Social Security Administration in the United States starting in 1977. The work on this project could easily be described as HCI work, including task analyses, scenario generation, screen prototyping, and building a usability laboratory. Pew also describes presenting some of his work at the annual meeting of the Human Factors Society in 1979. Ben Shneiderman published *Software Psychology*, considered one of the first books on the topic of HCI, in 1980. The terms “office automation” and “office information systems” were popular in the late 1970s. At that time, you could find articles that could be considered HCI-related, in fields such as management, psychology, software engineering, and human factors. In an interesting article on the history of office automation systems, Jonathan Grudin describes 1980 as the “banner year” for the study of office automation systems, after which, the number of people studying the topic dwindled, and many of them refocused under the title of HCI (Grudin, 2006b). The computer mouse was first publicly demoed by

Doug Engelbart in 1968 ([Engelbart, 2016](#)). Still others point to seminal papers as far back as Vannevar Bush's "As We May Think," which looks surprisingly relevant, even today ([Bush, 1945](#)).

In the late 1970s and early 1980s, computers were moving out of the research laboratory and "secure, cooled room" into the home and the office. The use of mainframes was transitioning into the use of mini- and then microcomputers, and the more popular personal computers were making their debut: Apple II series, IBM PC/XT, and the Commodore/Vic. It was this move, away from large computers in secure rooms used only by highly trained technical people, to personal computers on desktops and in home dens used by nontechnical people in much greater numbers that created the need for the field of HCI. Suddenly, people were using computers just as a tool to help them in their jobs, with limited training, and personal computers became a product marketed to home users, like stoves or vacuum cleaners. The interaction between the human and the computer was suddenly important. Nonengineers would be using computers and, if there wasn't a consideration of ease of use, even at a basic level, then these computers were doomed to failure and nonuse. In the current context, where everyone is using computers, that may sound a bit odd, but back in the 1970s, almost no one outside of computing, engineering, and mathematics specialists were using computers. Personal computers weren't in school classrooms, they weren't in homes, there were no bank cash machines, or airline self check-in machines, before this shift towards nonengineering use happened. This shift created a sudden need for the field of HCI, drawing on many different fields of study.

1.2 TYPES OF HCI RESEARCH CONTRIBUTIONS

The field of HCI draws on many different disciplines, including computer science, sociology, psychology, communication, human factors engineering, industrial engineering, rehabilitation engineering, and many others. The research methods may have originated in these other disciplines. However, they are modified for use in HCI. For instance, techniques such as experimental design and observation from psychology, have been modified for use in HCI research. Because HCI draws on the work in so many different disciplines, people often ask "what is considered HCI research? What types of effort are considered research contributions?" In a recent article that we believe will become a classic read, [Wobbrock and Kientz \(2016\)](#) discuss seven types of research contributions:

Empirical contributions—data (qualitative or quantitative) collected through any of the methods described in this book: experimental design, surveys, focus groups, time diaries, sensors and other automated means, ethnography, and other methods.

Artifact contributions—the design and development of new artifacts, including interfaces, toolkits, and architectures, mock-ups, and "envisionments." These artifacts, are often accompanied by empirical data about feedback or usage. This type of contribution is often known as HCI systems research, HCI interaction techniques, or HCI design prototypes.

Methodological contributions—new approaches that influence processes in research or practice, such as a new method, new application of a method, modification of a method, or a new metric or instrument for measurement.

Theoretical contributions—concepts and models which are vehicles for thought, which may be predictive or descriptive, such as a framework, a design space, or a conceptual model.

Dataset contributions—a contribution which provides a corpus for the benefit of the research community, including a repository, benchmark tasks, and actual data.

Survey contributions—a review and synthesis of work done in a specific area, to help identify trends and specific topics that need more work. This type of contribution can only occur after research in a certain area has existed for a few years so that there is sufficient work to analyze.

Opinion contributions—writings which seek to persuade the readers to change their minds, often utilizing portions of the other contributions listed above, not simply to inform, but to persuade.

The majority of HCI research falls into either empirical research or artifact contributions, and this book specifically addresses empirical research using all of the potential data collection methods utilized in empirical research. In their analysis of research papers submitted to the CHI 2016 conference, Wobbrock and Kientz found that paper authors indicated in the submission form that over 70% of the papers submitted were either empirical studies of system use or empirical studies of people, and 28.4% were artifact/system papers (it is important to note that authors could select more than one category, so percentages can add up to more than 100%). There were a fair number of papers submitted on methodological contributions, but submissions in all of the other categories of contributions were rare (Wobbrock and Kientz, 2016). This provides some empirical data for what we (as book authors) have observed, that most HCI research is either empirical or systems research (or sometimes, a combination of both, such as when you develop a prototype and have users evaluate it).

1.3 CHANGES IN TOPICS OF HCI RESEARCH OVER TIME

The original HCI research in the 1980s was often about how people interacted with simple (or not so simple) office automation programs, such as word processing, database, and statistical software. The basics of interfaces, such as dialog boxes, and error messages, were the focus of much research. Some of the classic HCI articles of the 1980s, such as Norman's analysis of human error (Norman, 1983), Carroll's "training wheels" approach to interface design (Carroll and Carrithers, 1984), and Shneiderman's work on direct manipulation (Shneiderman, 1983) are still very relevant today. Towards the late 1980s, graphical user interfaces started to take hold. In the late 1980s and early 1990s, there was growth in the area of usability engineering methods (and the Usability Professionals' Association, now known as UXPA, was founded in 1991). But there was a major shift in the field of HCI research during the early to mid 1990s, as the Internet and the web gained wide acceptance. New types of interfaces and communication, such as web pages, e-mail, instant messaging,

and groupware, received attention from the research community. This caused an increased number of research fields to be included under the umbrella of HCI, especially communication. A recent article by [Liu et al. \(2014\)](#) on trends of HCI research topics, determined a big difference between research in 1994–2003, which focused on fixed technology, and research from 2004–13, which focused on mobile and portable computing (such as tablets and smart phones).

Around 2004–05, the focus of research shifted more towards user-generated content that was shared, such as photos, videos, blogs, and wikis, and later grew into research on social networking. On Dec. 26, 2006, Time Magazine famously named “You” as the “person of the year” for generating much of the content on the web. The topic of user diversity gained more attention, with more research studying how younger users, older users, and users with disabilities, interact with technologies. In the late 2000s, research increased on touch screens, especially multitouch screens, with studies on motor movement focused on pointing using fingers, rather than computer mice. It is important to note that while multitouch screens only entered common public use in the late 2000s, multitouch screens had been developed and researched as far back as the early 1980s ([Buxton, 2016](#)).

The research focus in the late 2010s (the publication date of the book) is no longer on something as simple as task performance in statistical software, but is now more focused on collaboration, connections, emotion, and communication (although, again, research on collaboration has existed since the early 1980s, even if it's now just gaining attention). The focus is not just on workplace efficiency any more, but is on whether people like an interface and want to use it, and in what environment they will be using the technology. Today's research focuses on topics such as mobile devices, multitouch screens, gestures and natural computing, sensors, embedded and wearable computing, sustainability, big data, social and collaborative computing, accessibility, and other topics ([Liu et al., 2014](#)). But, of course, that will change over time! The topics of HCI research continue to change based on factors such as technological developments, societal needs, government funding priorities, and even user frustrations.

1.4 CHANGES IN HCI RESEARCH METHODS OVER TIME

There are many reasons why, over time, research methods naturally evolve and change. For instance, tools for research that were originally very expensive, such as eye-tracking, sensors, drones, facial electromyography (EMG), and electroencephalography (EEG) are now relatively inexpensive or at least are more reasonable, allowing more researchers to afford them and integrate these tools into their research. New tools develop over time, for instance, Amazon's Mechanical Turk. New opportunities present themselves, such as with social networking, where suddenly, there are billions of pieces of text and multimedia that can be evaluated, looking for patterns. Or with personal health tracking, or electronic health records, which allow for analysis of millions of data points, which have already been collected. Some types of research are now fully automated. For instance, years ago, researchers would do

a citation analysis to understand trends in research, but most of that analysis is now easily available using tools such as Google Scholar. On the other hand, automated tools for testing interface accessibility, are still imperfect and have not yet replaced the need for human evaluations (either with representative users or interface experts).

One important difference between HCI research and research in some of the other social sciences (such as sociology and economics), is that, large entities or government agencies collect, on an annual basis, national data sets, which are then open for researchers to analyze. For instance, in the United States, the General Social Survey, or government organizations such as the National Center on Health Statistics, the US Census Bureau, or the Bureau of Labor Statistics, collect data using strict and well-established methodological controls. Outside of the US, agencies such as Statistics Canada, and EuroStat, collect excellent quality data, allowing researchers to, in many cases, to focus less on data collection and more on data analysis. However, this practice of national and/or annual data sets, does not exist in the area of HCI. Most HCI researchers must collect their own data. So that alone makes HCI research complex.

Typically, HCI research has utilized smaller size datasets, due to the need for researchers to recruit their own participants and collect their own datasets. However, as the use of big data approaches (sensors, text analysis, combining datasets collected for other purposes) has recently increased, many researchers now utilize larger pools of participant data in their research. Whereas, studies involving participants might have had 50 or 100 users, it is common now to see data from 10,000–100,000 users. That is not to say that researchers have actually been interacting with all of those users (which would be logistically impossible), but data has been collected from these large data sets. Doing research involving 100,000 users versus 50 users provides an interesting contrast. Those 100,000 users may never interact with the researchers or even be aware that their data is being included in research (since the terms of service of a social networking service, fitness tracking, or other device, may allow for data collection). Also, those participants will never get to clarify the meaning of the data, and the researchers, having no opportunity to interact with participants, may find it hard to get a deeper understanding of the meaning of the data, from the participants themselves. Put another way, big data can help us determine correlations (where there are relationships), but might not help us determine causality (why there are relationships) (Lehikoinen and Koistinen, 2014). On the other hand, by interacting with participants in a smaller study of 50 participants, researchers may get a deeper understanding of the meaning of the data. Combining big data approaches with researcher interaction with a small sampling of users (through interviews or focus groups) can provide some of the benefits of both approaches to data collection, understanding not only the correlations, but also the causality (Lehikoinen and Koistinen, 2014).

Another important difference between HCI research and research in some of the other fields of study is that longitudinal studies in HCI are rare. Fields such as medicine may track health outcomes over a period of decades. National census data collection can occur over centuries. However, longitudinal data generally does not exist in the area of HCI. There could possibly be a number of reasons for this. Technology in general, and specific tools, change so rapidly that, a comparison of computer usage in 1990, or even

2000, versus 2017 might simply not be relevant. What would you compare? However, a trend analysis over time might be useful, because there are some audiences for HCI research, for whom trend analyses, over time, are considered a primary approach for data collection (such as the CSCW researchers described in [Section 1.6](#) and the policy-makers described in [Section 1.7](#)). Furthermore, there are areas of HCI research where longitudinal data would be both appropriate and very relevant. For instance, Kraut has examined, over a 15-year period, how internet usage impacts psychological well-being, and how the types of communication, and the trends, have changed over time ([Kraut and Burke, 2015](#)). There are other similar longitudinal studies that are also very useful, for instance, documenting that 65% of American adults use social networking tools in 2015, up from 7% in 2005 ([Perrin, 2015](#)), or documenting internet usage trends over a 15 year period ([Perrin and Duggan, 2015](#)). One could easily imagine other longitudinal studies that would be useful, such as how much “screen time” someone spends each day, over a 20 year period. The lack of longitudinal research studies in HCI, is a real shortcoming, and in some cases, limits the value that communities outside of computer science, place on our research.

Another reason why HCI research is complex is that, for much of the research, not just any human being is appropriate for taking part as a participant. For instance, a practice in many areas of research, is simply to recruit college students to participate in the research. This would certainly be appropriate if the focus of the research is on college students. Or this potentially *could* be appropriate if the focus of the research is on something like motor performance (in which the main factors are age and physiological factors). However, for much of HCI research, there is a focus on the users, tasks, and environments, which means that not only must the users be representative in terms of age, educational experience, and technical experience, but also in terms of the task domain (it is often said that you must “know thy user”). For instance, that means that to study interfaces designed for lawyers, you must actually have practicing lawyers taking part in the research. It will take time to recruit them, and they will need to be paid appropriately for their participation in a research study. Perhaps it is possible, although not ideal, to substitute law students in limited phases of the research, but you would still need to have actual practicing lawyers, with the right task domain knowledge, taking part in the research at the most critical phases. Recruitment of participants is much more complex than just “find some people,” and it can be actually quite complex and take a fair amount of time. For someone coming from a background of, say, sociology, the number of participants involved in HCI studies can seem small, and the focus may be different (strict random sampling in sociology, versus representativeness in HCI). But our goals are also different: in HCI, we are primarily trying to study interfaces, and how people interact with interfaces, we are not primarily studying people, so we don’t always necessarily have to claim representativeness.

Despite historic roots in the early 1980s, only in the last 10–15 years or so have individuals been able to graduate from universities with a degree that is titled “Human-Computer Interaction” (and the number of people with such a degree is still incredibly small). Many people in the field of HCI may have degrees in computer

science, information systems, psychology, sociology, or engineering. This means that these individuals come to the field with different approaches to research, with a certain view of the field. Even students studying HCI frequently take classes in psychology research methods or educational research methods. But taking just an educational or psychological approach to research methods doesn't cover the full breadth of potential research methods in HCI. Ben Shneiderman said that “The old computing is about what computers can do, the new computing is about what people can do” ([Shneiderman, 2002](#)). Since HCI focuses on what people can do, it involves multiple fields that involve the study of people, how they think and learn, how they communicate, and how physical objects are designed to meet their needs. Basically, HCI researchers need all of the research methods used in almost all of the social sciences, along with some engineering and medical research methods.

1.5 UNDERSTANDING HCI RESEARCH METHODS AND MEASUREMENT

HCI research requires both rigorous methods and relevance. It is often tempting to lean more heavily towards one or the other. Some other fields of research do focus more on theoretical results than on relevance. However, HCI research must be practical and relevant to people, organizations, or design. The research needs to be able to influence interface design, development processes, user training, public policy, or something else. Partially due to the philosophies of the founders of the field, HCI has had a historic focus on practical results that improve the quality of life ([Hochheiser and Lazar, 2007](#)). Is there a tension sometimes between researchers and practitioners? Absolutely. But all HCI research should at least consider the needs of both audiences. At the same time, the research methods used (regardless of the source discipline) must be rigorous and appropriate. It is not sufficient to develop a new computer interface without researching the need for the interface and without following up with user evaluations of that interface. HCI researchers are often placed in a position of evangelism where they must go out and convince others of the need for a focus on human users in computing. The only way to back up statements on the importance of users and human-centered design is with solid, rigorous research.

Due to this interdisciplinary focus and the historical development of the field, there are many different approaches to measurement and research currently used in the field of HCI. A group of researchers, all working on HCI-related topics, often disagree on what “real HCI research” means. There are major differences in how various leaders in the field perceive the existence of HCI. Be aware that, as an HCI researcher, you may run into people who don't like your research methods, are not comfortable with them, or simply come from a different research background and are unfamiliar with them. And that's OK. Think of it as another opportunity to be an HCI evangelist. (Note: As far as we know, the term “interface evangelist” was first used to describe Bruce Tognazzini. But we really think that the term applies to all of us who do HCI-related work.) Since the goal of this book is to provide a guide that

introduces the reader to the set of generally accepted empirical research practices within the field of HCI, a central question is, therefore, how do we carry out measurement in the field of HCI research? What do we measure?

In the early days of HCI research, measurement was based on standards for human performance from human factors and psychology. How fast could someone complete a task? How many tasks were completed successfully, and how many errors were made? These are still the basic foundations for measuring interface usability and are still relevant today. These metrics are very much based on a task-centered model, where specific tasks can be separated out, quantified, and measured. These metrics include task correctness, time performance, error rate, time to learn, retention over time, and user satisfaction (see [Chapters 5 and 10](#) for more information on measuring user satisfaction with surveys). These types of metrics are adopted by industry and standards-related organizations, such as the National Institute of Standards and Technology (in the United States) and the International Organization for Standardization (ISO). While these metrics are still often used and well-accepted, they are appropriate only in situations where the usage of computers can be broken down into specific tasks which themselves can be measured in a quantitative and discrete way.

Shneiderman has described the difference between micro-HCI and macro-HCI. The text in the previous paragraph, improving a user's experience using well-established metrics and techniques to improve task and time performance, could be considered micro-HCI ([Shneiderman, 2011](#)). However, many of the phenomena that interest researchers at a broader level, such as motivation, collaboration, social participation, trust, and empathy, perhaps having societal-level impacts, are not easy to measure using existing metrics or methods. Many of these phenomena cannot be measured in a laboratory setting using the human factors psychology model ([Obrenovic, 2014](#); [Shneiderman, 2008](#)). And the classic metrics for performance may not be as appropriate when the usage of a new technology is discretionary and about enjoyment, rather than task performance in a controlled work setting ([Grudin, 2006a](#)). After all, how do you measure enjoyment or emotional gain? How do you measure why individuals use computers when they don't have to? Job satisfaction? Feeling of community? Mission in life? Multimethod approaches, possibly involving case studies, observations, interviews, data logging, and other longitudinal techniques, may be most appropriate for understanding what makes these new socio-technical systems successful. As an example, the research area of Computer-Supported Cooperative Work (CSCW) highlights the sociological perspectives of computer usage more than the psychological perspectives, with a focus more on observation in the field, rather than controlled lab studies ([Bannon, 2011](#)).

The old methods of research and measurement are comfortable: hypothesis testing, statistical tests, control groups, and so on. They come from a proud history of scientific research, and they are easily understood across many different academic, scientific, and research communities. However, they alone are not sufficient approaches to measure all of today's phenomena. The same applies to the “old standard” measures of task correctness and time performance. Those metrics

may measure “how often?” or “how long?” but not “why?” However, they are still well-understood and well-accepted metrics, and they allow HCI researchers to communicate their results to other research communities where the cutting-edge tools and research methods may not be well-understood or well-accepted.

You may not be able to use experimental laboratory research to learn why people don't use technology. If you want to examine how people use portable or mobile technology such as smart phones and wearable computing, there are limitations to studying that in a controlled laboratory setting. If you want to study how people communicate with trusted partners, choose to perform business transactions with someone they don't know on another continent (as often happens with Ebay), or choose to collaborate, you need to find new ways of research and new forms of measurement. These are not research questions that can be answered with quantitative measurements in a short-term laboratory setting.

Consider Wikipedia, a collaborative, open-source encyclopedia. Currently, more than five million articles exist in English on Wikipedia, with an estimate of 70,000 active contributors (<https://www.wikipedia.org>), who spend their own time creating and editing Wikipedia entries. What causes them to do so? What do they get out of the experience? Clearly, task and time performance would not be appropriate metrics to use. But what metrics should be used? Joy? Emotion? A feeling of community? Lower blood pressure? This may not be a phenomenon that can be studied in a controlled laboratory setting (Menking and Erickson, 2015). The field of HCI has begun to apply more research methods from the social sciences, and we encourage the reader to start using some new research approaches that are not even in this textbook! Please be aware that people from other disciplines, as well as your “home discipline,” will probably challenge the appropriateness of those research methods!

1.6 THE NATURE OF INTERDISCIPLINARY RESEARCH IN HCI

Interdisciplinary research using multiple research methods, is not always easy to do. There are many challenges that can arise, in many cases due to the individual cultures of each of the disciplines involved. The HCI community might be considered by some to be an interdisciplinary community, a multidisciplinary community, or its own discipline (Blackwell, 2015). Regardless of the status of HCI as interdisciplinary, multidisciplinary, or its own discipline, many conferences, professional organizations, and academic departments keep the focus on their primary discipline. When interdisciplinary research gets filtered through single-discipline evaluations, there are many challenges that can occur. Some of the challenges are well-known, such as how some disciplines (e.g., computer science) focus more on conference publications and others (e.g., management information systems) focus on journal publications (Grudin, 2006a). Some disciplines focus on single-author publications, while others focus primarily on group-author publications. Some disciplines are very open about sharing their results, while others keep their results more confidential. Some disciplines are very self-reflective and do research studies about their discipline

(trends of research, rankings, funding, collaborations), while others do not. Some disciplines are primarily focused on getting grant money, while other disciplines are less interested, or can even be leery of the influences of outside sponsors. Even the appropriate dress at conferences for each discipline can vary widely. It is important, for a number of reasons, to become familiar with the research methods and preferences in different disciplines. You need to be able to communicate your research methods, and the reasons why you chose some and not others, in a very convincing way. When you submit journal articles, conference papers, grant proposals, or book chapters, you never know who will be reviewing your work. The chances are good that your work will be reviewed by people who come from very different research backgrounds, and interdisciplinary researchers can sometimes have problems convincing others at their workplace of the quality and seriousness of their work. But all of these are primarily concerns with an individual's professional career or with administrative issues (Sears et al., 2008).

There are more serious, but less well-known, challenges related to interdisciplinary research. As discussed earlier in this chapter, no research method, approach, or discipline is perfect. A research project is a series of steps and decisions related to data collection. For instance, there is a theoretical foundation for the data collection effort, there is a research method involved, often human participants are recruited and involved, there is data analysis, and then there is the discussion of implications involved. The development of a proof-of-concept or prototype is also frequently involved. Depending on the majority disciplinary background of those involved in the research, there may be different perspectives, value systems, and expectations (Hudson and Mankoff, 2014). For instance, there could be a distinction between technical HCI research (focused on interface building) versus behavioral HCI research (focused on cognitive foundations) which would likely have different expectations in terms of number and background of participants, development of a tool or interface, and outcomes (Hudson and Mankoff, 2014).

Different disciplines can sometimes be most interested in, and more focused on, different steps in the research process. While no one would ever say, "I'm not interested in the research methods," in many cases, there are steps that are considered to be of less interest to people from a certain discipline. And there may be historical roots for that. For instance, as described earlier and in other chapters, there are large data collection efforts that use strict controls, in fields such as sociology, and those data sets are available for researchers internationally to analyze. However, as previously discussed, no such central data sets exist for HCI and it is not considered a standard practice to publish your data sets or make them available to others. It is a very different model in other fields. That may lead to a focus on certain stages of research more than others.

(Please note: we expect the following paragraphs to be a bit controversial; however, we do believe strongly, based on our experience, that they are true.) One discipline might have an expectation that a specific step (such as research design) is done "perfectly," but that it is acceptable to give more flexibility in other steps (such as the types of participants). The management information systems community of

HCI researchers has a well-known focus on the theoretical underpinnings of any research. Computer science-based HCI researchers often have less interest in theory and much more of an interest in the practical outcomes of the research on interfaces (although [Carroll, 2003](#) is a noteworthy effort on theory in HCI). This distinction is seen, for instance, in the Technology Acceptance Model, which is core theory and has central importance for HCI researchers focused on management information systems ([Davis, 1989](#); [Venkatesh and Davis, 2000](#)), but is not well-known to the HCI researchers focused on computer science. While general computer science researchers have a great deal of theory in, say, algorithms, HCI research in computer science does not have a major focus on theory.

When having interdisciplinary discussions and working on interdisciplinary teams, it's important to be aware of these distinctions. Sociology-based HCI research tends to focus on the demographics of the research participants and determining if they are a true random sample, while this is not considered critical in computer science, where computer science students are often used as participants (even when it is not appropriate). Psychology-based HCI research tends to focus on an ideal and clean research design. HCI research based on computer science and on design is focused more on the implications for interfaces, although computer science may focus more on the technical underpinnings while design focuses more on the look and feel of the interface. These are just generalizations, obviously; all disciplines want excellence at all stages of research, but it is true that disciplines tend to focus more intensely on particular stages of research. The good news is that we want all of these different groups focusing on improving each stage of the research process. We WANT different groups looking at research through their different lenses. We want to get that triangulation (described more in [Section 1.8](#)), where people look at the same research questions, using different methods, different approaches, and different lenses, over time, with the goal of discovering some scientific truths.

1.7 WHO IS THE AUDIENCE FOR YOUR RESEARCH?

Most researchers in HCI often, unknowingly, target their HCI research towards other researchers. The metrics that are used most often to ascertain impact of a research publication, relate to the number of times that a paper is cited in other publications, and impact factor of the journal or conference proceeding. Metrics used in many areas of science, such as the h-index, can be used to ascertain productivity of an individual researcher, rather than a specific article, but again, it is based primarily on how the specific researcher has impacted other researchers. Alternative metrics, such as tracking number of downloads, using microblogging (e.g., Twitter), online reference managers (e.g., Zotero and Mendeley) and blogging to track impact, are also gaining in popularity ([Bornmann, 2015](#)). However, these metrics are reflections of how a research publication impacts other researchers, not how a research publication has impact outside of the research world. The idea of societal impact outside of other publications, is not something that most researchers receive training on, or

even consider, and unless an individual is working in an industrial research lab or as a practitioner (where the goal is often to influence design and development) it is just perceived that the goal is to be cited by other researchers. However, there are other audiences for HCI research, aside from other researchers. Doing research targeted at other audiences requires different approaches to research, and different ways of communicating the findings of that research.

Outside of HCI researchers, the other audience that most HCI researchers would be familiar with, is the audience of individuals who do systems development and interface design, as practitioners. Often, industrial HCI labs focus on HCI systems research, with the goals of doing good, publishable research while testing out designs and/or influencing the next generation of product interfaces at the company or organization. Researchers at universities, may also partner with industry, to influence the interaction design in corporations or nonprofit organizations. Unlike HCI research aimed at researchers taking place in a university setting without industrial partners, there may be issues about disclosure, about sharing results publicly, about corporate secrecy. There also may be much more concern about the control of intellectual property resulting from the research.

Furthermore, the types of controls or inclusion criteria used in HCI research targeted at industrial impact, may differ from the types of controls utilized in HCI research targeted at other researchers. For instance, it can be expected that a company would be most interested in evaluating aspects related to their own products. So, when doing research to impact design and development, the company might only be interested in their own products and the specific configurations that the product is designed to work with. As an example, a company, researching how their new software application might be utilized for blind people, might only test it on certain operating systems (e.g., iOS only, rather than Windows, or only Windows 8 and later), or with certain screen readers (e.g., JAWS 14 or later, or Window-Eyes, but not VoiceOver). The product being evaluated by users, may have a specific configuration that it is designed to work with, and so the research may need to be limited to that configuration, even if that is an unrealistic configuration. For instance, a configuration may be unrealistic, either because no one is currently using that configuration, or because the configuration would bias the research since it would only allow for very advanced users who are on the cutting edge. Companies often face this challenge—there is a large installed base of users who utilize old versions of software or operating systems, yet this is not represented in the user research which involves only advanced users utilizing only the newest technologies, a situation that is not very representative.

Another potential target audience for HCI research is policymakers. Public policymakers need to have data to inform their decisions related to HCI issues, in the areas of statutory laws, regulations, executive orders, and everything from legal cases to human rights documents such as treaties. While many areas of science and technology have well-developed policy outreach, such community infrastructure does not yet exist for public policy issues related to HCI. There are a number of areas, where in the past, individual HCI researchers have been successful in informing and guiding public policy, and these include accessibility and ergonomics (Lazar et al., 2016). Furthermore, individuals from the HCI community have taken leadership roles as

government policymakers in countries such as the United States and Sweden. Many more areas exist where public policies have been created that influence HCI research work, often (and unfortunately) without the benefit of feedback from the HCI community. These areas where HCI research has been impacted include laws and regulations for human subjects research, standards for measurement, areas of research funding, language requirements for interface design, data privacy laws, and specific domains such as e-government, education, libraries, voting, and healthcare (Lazar et al., 2016). Because there is not an existing lobbying infrastructure, or entrenched interests on most HCI-related topics, this is a great opportunity for HCI researchers to have a true impact on public policies. Furthermore, some governments have legal limitations on how much data can be collected from citizens, so research studies (even a usability test involving 25 users) can be logistically hard for government to implement or even get approval for. However, the requirements of a university Institutional Review Board are often easier, and therefore, HCI researchers can often do data collection to inform policymakers, that a government agency may simply not be allowed to do.

When trying to perform HCI research with the target audience of public policymakers, there are some logistical considerations to be aware of. Policymakers in general, are very concerned with the number of people who are impacted (e.g., how many children or people with disabilities are within their representation area), and which specific laws or policies relate to your HCI work. So, computer scientists tend to make generalizations about items outside of computer science (e.g., “there is a law” or “lots of people”) but research targeted towards policymakers needs to be much more specific in terms of coverage. In general, policymakers like longitudinal research, because they like to know the trends in how people are being affected (e.g., is the situation getting better or worse?). Furthermore, it is important to understand the timelines of policymakers (e.g., when public comments are due on a regulatory process, when legislation is being considered, when legal proceedings occur), because, unlike in academia where there is always another conference to submit to, or another journal to submit your research to, when dealing with the timelines of policymakers, often there is no flexibility and if you miss a deadline, you will have zero impact (Lazar, 2014). Policymakers are not likely to communicate in the same way as researchers, so if you think that you can have an impact by just emailing or skyping with a policymaker, or sending them your research paper, you are mistaken. Policymakers tend to work only via face-to-face contact, so if you want to build relationships with policymakers, you need to schedule an appointment to meet with them. You also would be wise to provide summaries of research, designed for people who do not have a background in your area of HCI research (Lazar, 2014).

1.8 UNDERSTANDING ONE RESEARCH PROJECT IN THE CONTEXT OF RELATED RESEARCH

There is no such thing as a perfect data collection method or a perfect data collection effort. All methods, all approaches, all projects have a flaw or two. One data collection effort does not lead to a definitive answer on a question of research. In scientific

communities, the goal is generally for multiple teams to examine the same research question from multiple angles over time. Research results should be reported, with enough detail so that other teams can attempt to replicate the findings and expand upon them. Replication is considered an important part of validating research findings, even though it is rare in HCI and often gets very little attention ([Hornbaek et al., 2014](#)) (and many other fields of research have similar complaints). All of these efforts, if they come up with the same general findings over time, give evidence for the scientific truth of the findings. This is often known as “triangulation.” One data collection effort, yielding one paper, is interesting in itself but does not prove anything. If you have 15 teams of researchers, looking at similar research questions, over a period of 10 years, using multiple research methods, and they all come to the same general conclusion about a phenomenon, then there is some scientific proof for the phenomenon. The proof is even stronger when multiple research methods have been used in data collection. If all of the research teams replicate the exact same research methods over 10 years, then there is the remote possibility that the methods themselves are flawed. However, the weight of evidence is strengthened when multiple research methods are used.

Researchers often speak of a “research life cycle,” describing the specific steps in a research project. Depending on who you ask, the steps can differ: for instance, (1) designing research, (2) running data collection, and (3) reporting research ([Hornbaek, 2011](#)).

But there is another type of life cycle to consider: when you are entering a new area or subspecialty of research, which methods are likely to be utilized first? On the other hand, which methods require first having additional research in place? For instance, two of the three coauthors of this book have been involved with performing research to understand how people with Down syndrome (both children and adults) utilize technology and what their interface needs are. When we decided to do this research, there was no existing HCI research on people with Down syndrome. There was no base of literature to draw from. So we first started with an exploratory survey to understand how children and young adults utilize technology. Then we did a series of observations of adults with Down syndrome who were expert users about what their skills were, and how they gained those skills. Then we utilized a usability testing methodology to understand how adults with Down syndrome utilize social networking and touch screens. Once we had a base of understanding about the research topic with those three studies, only then did we do an experimental design (to understand the effectiveness of different authentication methods for people with Down syndrome). It would have been too premature to start with an experimental design method first, when so little was known about the population of users and how they interact with technology. The controls necessary for an experimental design, would have not yet been understood, so there would have been lots of phenomenon that were unknown and not controlled for. Often, when a research topic is new, it is important to start with a research method that can be utilized in a more exploratory way—such as surveys, interviews, focus groups, and ethnography. Then, with a basis of understanding from a few exploratory studies, research studies utilizing more structured research methods—such as experimental design, automated data collection, and time diaries, could be performed. That's not to say that such an order must occur—but such an order often does occur, because more

background research, more structure, is required for certain types of research methods. Shneiderman describes this as a three-step process: observation, intervention, and controlled experimentation. The understanding through the exploratory research, can be utilized to build prototypes or hypotheses for experimental design (Shneiderman, 2016).

Another aspect of the research life cycle is determining when controlled, in-laboratory studies should occur, versus studies “in the wild” (also known as field studies or in-situ studies). There is a great discussion in the research community about when each approach is situationally appropriate. For instance, some authors argue that field studies are most appropriate for mobile device research, since mobile devices are utilized in the field, with weather, noise, motion, and competing cognitive demands playing an important role in usage (Kjeldskov and Skov, 2014). Controlled environments and precise measurement may simply not be realistic for the usage of certain types of technologies, such as mobile devices. Another argument for the increased use of field studies, is that, as researchers come to understand more about what specific aspects of design lead to increased usability, then the next step is to understand how those technologies fit into the complex work, leisure, and family lives of individuals (Kjeldskov and Skov, 2014). Field studies may present interesting challenges related to informed consent, since the period of data collection, and who participates, in a controlled environment, may be easy to ascertain. But for example, data collection in a public space (in the wild), such as marathon or a rock concert, may pose questions about the inclusion of data from people who are not aware of the data collection and did not consent to participate (Anstead et al., 2014). One can imagine multiple approaches for which research methods to utilize and in what order (as described in previous paragraphs). So perhaps researchers might first do exploratory research in the wild, before moving to more controlled laboratory settings. Or perhaps researchers might first do controlled laboratory experiments, and then move their research into the wild and do field studies. There is not one answer that is right or wrong.

From personal experience, the authors can verify that both approaches are useful, and the combination of controlled studies and field studies, often gives you interesting findings that make you rethink your approaches. For instance, from the authors of this textbook, there were three research studies of a web-based security prototype, in a combination of controlled settings (university lab, workplace, home, and always on a consistent laptop), from three different groups of users, where the average task performance rate on a specific prototype was always over 90%. When that same web-based security prototype was placed on the web, with a much more diverse set of users utilizing the prototype, generally with a lower level of technical experience, and with technical environment being another factor (older browsers, slow download speeds, etc.), the average task performance rate was under 50%, a significant drop. No research method is ever perfect, and trying out different research methods to investigate similar phenomenon, helps you to more fully understand your area of study. It is important to note that an individual's viewpoint on controlled laboratory experiments versus field studies, may also be influenced by their individual disciplinary background, so, for instance, those with engineering backgrounds may lean more naturally towards laboratory experiments compared to those with an anthropology background.

In HCI, there are some situations where the evidence over time supports a specific finding. One clear example is the preference for broad, shallow tree structures in menu design (see the “Depth vs Breadth in Menus” sidebar). Multiple research studies have documented that broad, shallow tree structures are superior (in terms of user performance) to narrow, deep tree structures.

DEPTH VS BREADTH IN MENUS

Multiple research studies by different research teams, throughout the history of the HCI field, have examined the issue of the trade-off between depth and breadth in menus. Generally, tree structures in menu design can be implemented as narrow and deep (where there are fewer choices per level but more levels) or as broad and shallow (where there are more choices per level but fewer levels).

Figure 1.1 shows three menu structures.

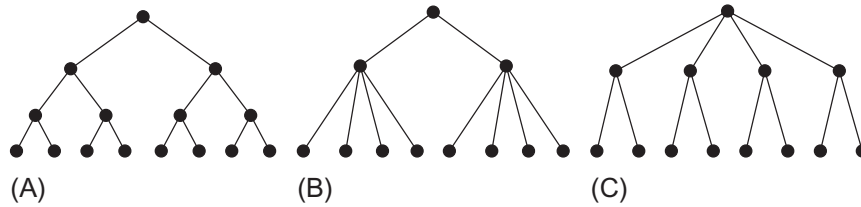


FIGURE 1.1

Types of tree structure in menu design: (A) narrow-deep: three levels with two choices at each level, (B) broad-shallow: two choices followed by four choices, (C) broad-shallow: four choices followed by two choices.

The research has consistently pointed to broad, shallow tree structures as being superior to narrow, deep structures. There are many possible reasons: users get more frustrated and more lost, the more levels they must navigate; users are capable of dealing with more than the 7 ± 2 options often cited in the research literature (since menus deal with recognition, not recall), and strategies for scanning can lead to superior performance. Different research methods and different research teams, examining different users, have all come to the same conclusion. So over time, the superiority of broad, shallow tree structures has become well-accepted as a foundation of interface design. Some of the better-known articles on this topic include:

- Hochheiser, H., Lazar, J., 2010. Revisiting breadth vs. depth in menu structures for blind users of screen readers. *Interacting with Computers* 22 (5), 389–398.
- Kiger, J.I., 1984. The depth/breadth trade-off in the design of menu-driven user interfaces. *International Journal of Man-Machine Studies* 20 (2), 201–213.
- Landauer, T.K., Nachbar, D.W., 1985. Selection from alphabetic and numeric menu trees using a touch screen: breadth, depth, and width.

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–78.

Larson, K., Czerwinski, M., 1998. Web page design: implications of memory, structure and scent for information retrieval. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 25–32.

Miller, D., 1981. The depth/breadth tradeoff in hierarchical computer menus. Proceedings of the Human Factors Society 25th Annual Meeting, pp. 296–300.

Snowberry, K., Parkinson, S., Sisson, N., 1983. Computer display menus. *Ergonomics* 6 (7), 699–712.

Wallace, D.F., Anderson, N.S., Shneiderman, B., 1987. Time stress effects on two menu selection systems. Proceedings of the Human Factors and Ergonomics Society 31st Annual Meeting, pp. 727–731.

Zaphiris, P., Mtei, L., 2000. Depth vs breadth in the arrangement of web links. Proceedings of the Human Factors and Ergonomics Society, 44th Annual Meeting, pp. 139–144.

In contrast to the example in the sidebar, other research topics in HCI still have no clear answer, with multiple studies that yield conflicting findings. For instance, what is the minimum number of people required for usability testing? See [Chapter 10](#), where the debate still rages on, as there is no agreed answer. The commonly repeated number is that 5 users is sufficient (although the research really doesn't say this), and more recent studies have suggested 10 ± 2 users ([Hwang and Salvendy, 2010](#)) or even more than 10 users ([Schmettow, 2012](#)). We suggest that readers turn to [Chapter 10](#) to continue this debate. There may also be some research questions to which the answers change over time. For instance, in the late 1990s, web users tended to find download speed to be one of the biggest frustrations ([Lightner et al., 1996](#); [Pitkow and Kehoe, 1996](#)). User habits and preferences are fluid and there may be changes over, say, a 20-year period (factors such as increased availability of broadband Internet access may also play a role). The biggest frustration for web users right now would most likely be viruses or spam. When the web first became popular in the mid-1990s, web-wide subject lists and in-site navigation were popular methods for finding items; now, search boxes are far more popular methods for finding what you want (and it is possible that the introduction of Google played a role). When it comes to user preferences, there can be many different influences, and these preferences may change over time. This is yet another reason why one research project, at one point in time, does not make a scientific fact.

You should never get disappointed or upset when you find out that another research team is working on a similar research question. You should get excited, because it means that both research teams are moving closer to the end goal of some definitive scientific answers. The chances are very high that your research method won't be exactly the same, your research questions won't be exactly the same, and your human participants won't be exactly the same. The fact that other research teams are interested

in this topic shows the importance of the research area and strengthens your findings. Perhaps you should be more worried if no one else is interested in your research.

1.9 INHERENT TRADE-OFFS IN HCI

It would at first seem that, with enough research, you could simply decide which design is best by optimizing some specific measurement, such as task performance or time performance. First of all, as discussed earlier in this chapter, socio-technical systems can rarely be reduced to two or three measurements, and there are many factors to be controlled for. We can do comparison studies of small differences in menu structure or some detailed aspect of interface design, but it is much harder to compare fundamental recastings of tasks. In addition, there are inherent conflicts in HCI research and design. We make trade-offs and accept “better solutions” rather than optimal solutions. We have multiple stakeholders and not all of them can be satisfied. Design is not simple and it's not an optimization problem. Good HCI research allows us to understand the various factors at play, which design features may work well for which users, and where there are potential conflicts or trade-offs.

For example, we can learn how to make interfaces that are far better than our current interfaces. However, users may not prefer those interfaces because they are so different from the current interfaces. So maybe we should modify our interfaces gradually, making only minor changes each time? Keyboards are a perfect example of this. We know how to make keyboards that are more ergonomic, with key layouts that allow for much faster typing. However, the keyboard layout predominantly used with the Roman alphabet is still the QWERTY key layout. Why? We have far superior designs. However, people have been comfortable with the QWERTY layout for years and the other key layouts have not caught on (despite their clear superiority from a design and usability point of view). So we still use the QWERTY layout. It's a trade-off. You want to make interfaces that are much better but users want consistency. In the short-term, a totally new interface lowers user performance, increases user error, and lowers user satisfaction. In the long-term, a modified interface may improve performance and result in higher satisfaction. This focus on very minor tweaks can be seen in the attention currently being paid, in industry and government, to the idea of A/B testing, where you test very minor interface changes, barely noticeable by the user, and then roll out those that are deemed to be successful, increasing traffic, increasing sales, and reducing costs ([Wolfers, 2015](#)). Of course, there are sometimes new interfaces, new devices, that just leap ahead with a totally different design and users love it, such as the Apple iPad tablet device. You shouldn't create a totally new design, apparently, unless it's something so cool that users want to spend the time to learn how to use it. Well, how do you measure that? How do you decide that? How do you plan for that? It's not easy.

Other examples of trade-offs in HCI also exist. For instance, the intersection of usability and security ([Bardram, 2005](#); [DeWitt and Kuljis, 2006](#)). In HCI, we want interfaces that are 100% easy to use. People focused on computer security want

computers that are 100% secure. By definition, many security features are designed to present a roadblock, to make users stop and think, to be hard. They are designed so that users may not be successful all of the time. The best way to make a 100% usable interface would be to remove all security features. Clearly, we can't do that. From the HCI point of view, our goal is to reduce unnecessary difficulty. Right now, the typical user has so many passwords that they simply can't remember them or they choose easy-to-remember (and easy to crack) passwords (Chiasson et al., 2008). Users may write their passwords on a sheet of paper kept in their wallet, purse, or desk drawer (none of which are secure), or they click on the feature that most web sites have saying, "Can't remember your password? Click here!" and their password is e-mailed to them (also not secure!). We suggest to readers to check out the annual ACM Symposium on Usable Privacy and Security (SOUPS) for research on the intersection of usability and security. Other inherent trade-offs occur in the area of sustainability. While people working in the field of information technology may often be focused on new and better devices and design, faster machines, and faster processing, this can lead to high energy usage and a lot of waste. Sustainability means trying to encourage users to limit their energy usage (Chetty et al., 2009), to keep using current devices, and to reduce the amount of technology waste by allowing current devices to be repaired or retrofitted, rather than just throwing the device out (Mankoff et al., 2007a). Millions of current personal computers end up in landfills, poisoning the earth and water.

Being user centered, as HCI tends to be, also means being concerned about the impacts of technology on human life. In the past, this meant that HCI researchers were interested in reducing repetitive strain injuries from computer usage, whether spending lots of time on the Internet made you depressed, and whether computer frustration could impact on your health. How does all of our technology creation, usage, and disposal impact on the quality of our life and the lives of future generations? Can persuasive devices and social networking be used to encourage us to lower our ecological footprint? (Gustafsson and Gyllenswärd, 2005; Mankoff et al., 2007b). Let's go back to our keyboard example: if all keyboards in the English-speaking world were changed over to a different key layout (say, the DVORAK layout), there might be some initial resistance by users but, eventually, user performance might improve. However, how would those millions of keyboards in landfill impact on the quality of human life? This is a new point to evaluate when considering how we do research in HCI. What is the ecological impact of our research? What is the ecological impact of new interfaces or devices that we build? While it is likely that we won't know in advance what type of ecological impact our research work will lead to, it's an important consideration as we do our research, yet another inherent challenge in HCI.

1.10 SUMMARY OF CHAPTERS

Given that the topic of research methods in HCI is so broad, we have tried to give approximately one chapter to each research method. However, the book starts out with three chapters revolving around the topic of experimental design. Whole books