

RESEARCH METHODS IN HUMAN-COMPUTER INTERACTION

Jonathan Lazar
Jinjuan Heidi Feng
and
Harry Hochheiser

 **WILEY**

A John Wiley and Sons, Ltd., Publication

This edition first published 2010
© 2010 John Wiley & Sons Ltd

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

Reprinted with corrections May 2010

Reprinted September 2010

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Lazar, Jonathan.

Research methods in human-computer interaction / Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-72337-1 (pbk.)

I. Human-computer interaction—Research. I. Feng, Jinjuan Heidi. II. Hochheiser, Harry. III. Title.

QA76.9.H85.L396 2010

004.01'9—dc22

2009033762

A catalogue record for this book is available from the British Library.

Typeset in 11/13pt Bembo by Aptara Inc., New Delhi, India.

Printed in Great Britain by Bell & Bain, Glasgow.

14

Working with human subjects

14.1 Identifying potential participants

14.2 Care and handling of research participants

14.3 Online research

Research into human-computer interaction (HCI) almost invariably involves the participation of human subjects. Whether you are running a focus group, leading a collaborative design process, running a controlled study, or conducting an ethnographic investigation, you need to engage people in your work.

Although this may sound simple, it isn't. As anyone who has done so can tell you, working with human subjects involves many challenges. Finding the right subjects is often difficult and time-consuming, especially for evaluations of systems designed for specific populations or situations.

The real fun can begin when the subjects are ready to begin participating in your study. Research ethics require that participants must be treated fairly and with respect. This means that they must be provided with information about the nature of the study that they can use to make a meaningful decision as to whether or not they really want to be involved. This notion of *informed consent* is a critical component of modern research on human subjects.

Although some of the details may differ, the general problems in finding and informing research subjects apply to any form of research involving human participants, regardless of the type of person involved. The additional challenges that online research presents in each of these areas is described in Section 14.4. This chapter uses the terms "subject" and "participant" interchangeably.

14.1 Identifying potential participants

You've just built an interactive system for two-handed input to an architectural modeling system and you'd like to run some summative tests to help find usability problems. This leaves you with a problem: who should participate in your study? There are plenty of potential users with two hands, but having the physical ability to manipulate your tool is just a start. People without the appropriate training and experience will be unable to tell you if your tool succeeds in its primary goal – supporting the work of an architect. Narrowing your pool of potential participants to architects would be your next logical step, but even this limitation may not be fine-grained enough. Are you willing to accept architecture students? This might help if there is a school of architecture nearby, but students may lack real-world experience. This might lead you to insist upon professional architects, who may be hard to find. HCI researchers are familiar with these and related challenges in finding appropriate study participants.

In the early days, in the late 1970s and early 1980s, many of the participants in HCI research were workers in corporate computing environments. This population of relatively early users was professionally motivated to participate in studies aimed at improving the systems that they used. As computer use spread more broadly into society and academic groups became active centers of HCI research, student bodies became available (often just walking down the hall) and easily motivated (via cash or food) pools of participants. Countless studies involving computer science or psychology undergraduates have been published over the years.

So, what's wrong with recruiting undergraduate students – or other easily found subjects – in HCI research? Often, nothing. If you are interested in evaluating interfaces intended for use by undergraduate students, this approach is perfect. However, tests that draw on a homogeneous group of participants may be open to criticism: results may not apply to users from a different demographic group. Even if a specific menu arrangement in a word-processing program works well for (predominantly young, male) computer science students, it may not work well for retired women. In a case like this, the mismatch may simply limit the extent to which you can claim that your study answers the problem.

The number of participants is another crucial factor. Different forms of research require different numbers of participants. Studies with too few participants may not yield generalizable results, while studies with too many participants are unnecessarily expensive and time-consuming.

14.1.1 Which subjects?

In selecting participants, you should strive to find people with *personal attributes and goals* appropriate for your study. By personal attributes, we mean demographic, educational, vocational and avocational details. Some studies may simply need computer users, while others need participants of a certain gender, age range, education level, professional background, or any combination of these characteristics.

Each individual's goals, background, and motivations may play a role in determining how appropriate they are for your study. Insufficiently interested subjects may be unlikely to contribute constructively, no matter how well they match your other criteria. Even with the right physical attributes, an architect who is strongly opposed to the use of computers for modeling would probably not make a good subject for studying the architectural tool described above. On the other hand, some studies might benefit from the perspective of less-motivated participants, who might be more critical and less forgiving of shortcomings than enthusiasts. The participation of these less-motivated users can be particularly useful when studying tools that may be used by a broad range of users in non-voluntary circumstances, such as mandatory workplace timesheet reports. Unmotivated users can also be useful for studies aimed at understanding the factors that might influence reluctance to adopt new technology.

Expertise is always an important consideration: study participants should have expertise that is comparable to that of the expected users. We usually define expertise in terms of two largely separable dimensions: computer expertise and domain expertise – knowledge of the problems, systems, goals, and tools used in a specific line of work. If you are testing a tool that is built for highly trained professionals who rarely use complex computer applications, you'll be looking for users who may be computer novices, even though they have significant domain expertise. In other cases, you might be looking for sophisticated computer users who are using a new type of software: computer experts but domain novices. Any differences in expertise between your target population and the participants in your study may lead

to results that are hard to interpret. You may be left wondering why your population of computer experts failed to successfully complete tasks with your interface for domain experts: was it because the interface failed or because the users lacked the required experience in the domain?

Interfaces that are intended for use by a broad audience present relatively little difficulty in terms of user characteristics. General-purpose desktop computing tools and interfaces on widely used communications devices are likely to be used by many motivated users, so study participants do not need to meet many specific criteria and can often (but not always) be similar to each other.

The need for appropriate participants becomes more apparent with tools that are designed for specific populations. Children and adults have vastly different cognitive and physical abilities, which directly influence their ability to act as useful study participants. Similarly, cultural differences between users may play a significant role in task performance for communication systems. Whenever possible, studies of tools designed for specific ages, genders, social backgrounds, and physical or cognitive abilities should involve participants who fit the appropriate category. Ethnographic studies of specific users and situations are also sensitive to the appropriateness of the participants. If study participants are not the intended users of a system, you can only make limited claims about the utility of the system for the intended population.

Systems designed for domain experts can be particularly challenging in this regard. As the construction of tools for highly specialized tasks requires a detailed understanding of domain-specific work practices, there is a natural tendency to use techniques such as participatory design to involve users in system design. This inclusion may lead to valuable insights, but domain experts who were involved in the design of a tool may have biases in favor of the resulting design, making them inappropriate candidates for subsequent usability tests or other summative evaluations.

Difference between users can also be an important part of study design. Investigations of potential gender differences in organizing certain forms of information would require both male and female participants. Similarly, an experiment exploring the role of user motivation in understanding the effectiveness of a given interface design may need participants who are highly motivated, as well as those who are not at all motivated.

Additional care is necessary when study designs require multiple groups that differ in some dimension. Ideally, the groups would differ in the relevant attribute but be comparable in all others. Any other differences would be possible confounding variables – factors that could be responsible for observed differences. In the study of gender differences in information management, the male and female groups should be comparable in terms of education, age, income, professional experience, and as many other factors as possible. If the women were significantly younger than the men, it might be hard to determine whether any performance differences were due to age or gender: further experimentation may be necessary.

Although these issues may be most important for controlled experiments, the identification of an appropriately general group of participants is always a challenge. Appropriate recruiting methods can help, but there are no guarantees. Despite your best efforts to find a representative population, you always face the possibility that your group of participants is insufficiently representative in a way that was unanticipated. As this bias is always possible, it's best to explicitly state what steps you have taken to account for potentially confounding variables and to be cautious when making claims about your results.

14.1.2 How many subjects?

Determining the number of participants to involve in a research study is subject to a trade-off between the information gained in the study and the cost of conducting it. Studies with a very large number of participants – say, tens of thousands – probably involve many people of different ages, educational backgrounds, and computer experience. Any outcome that you see consistently from this population may therefore not be something that can be explained away by the specific characteristics of the individual participants: it is likely to be a “real” effect. Huge studies like this are particularly helpful for controlled experiments in search of statistically significant results. Even subtle differences can be statistically significant if the populations are sufficiently large.

Unfortunately, large studies are difficult and expensive to run. Each participant involves substantial costs for recruiting, enrolling, conducting the study, and managing data. If the participants are not at your workplace, there may be travel involved, and many studies pay people for their time. If your study allows you to involve many people at once – perhaps 20 people in a roomful of computers – you may be able to achieve some efficiencies in terms of the time involved. However, research that involves one-on-one interactions between a researcher and a participant may have costs that grow linearly with the number of participants.

At the other extreme, a study with one individual has very real limitations. This study would be relatively inexpensive, but also very limited. Because this study would not have a range of users with different characteristics, any results would run the risk of telling you more about the participant than they did about the research question at hand. If you're conducting an ethnographic study with one person, you may learn a great deal about how that person performs certain types of work, but you have no idea about how representative her habits are: you may get unlucky and find someone who is completely unlike colleagues in the field. As studies with few participants rarely, if ever, produce statistically significant results, the conclusions that you can draw from these small studies are extremely limited.

Controlled experiments or empirical studies require a sample group of participants large enough to produce statistically significant results. The research design (the number of independent variables, within or between subjects) will play a role as well. Experiments involving larger numbers of independent variables and between-subjects (as opposed to within-subjects) analysis can require more participants (see Chapter 3). Limitations on resources can often lead researchers to substitute the feasible experiment – the design that

requires fewer participants – for the experiment they'd prefer to be doing. In some cases, statistical techniques can be used to determine the minimum number of subjects necessary for a result of a given significance (Chapter 3). Usually, you want at least 15–20 participants: smaller studies may miss potentially interesting results.

The inclusion of more participants gives you more statistical power. As each participant comes with costs in time, energy, and money, there are always good arguments in favor of limiting the size of the study. However, larger populations – ranging from several dozen to several hundred participants – offer the possibility of stronger statistical significance or the identification of subtle effects that would not be significant in smaller populations.

Statisticians have developed a range of techniques for determining the number of participants necessary for establishing statistically significant effects with differing degrees of confidence: Cook and Campbell, (1979) is a classic text in this area. These techniques can help you understand how many participants you need *before* your study starts, thus minimizing the chances for painful problems further down the line.

By contrast, case studies and ethnographic studies (Chapters 7 and 9) can often be conducted with a small number of users. If your goal is to gather requirements from domain experts, in-depth discussions with two or three motivated individuals may provide a wealth of data. The length of the session also plays a role here: ethnographic observations generally take more time per participant – and therefore place more demands upon the participants – than controlled experiments.

Usability studies can also be successfully conducted with a small set of participants. These studies use guidelines, heuristics, and a variety of techniques to identify potential usability problems with proposed interface designs (Chapter 10). While some authors state that as few as five usability experts could find 80% of the usability problems in an interface, there is a healthy debate about this (Nielsen, 1994). Of course, identification of an appropriate set of five usability experts might be a challenge, particularly since colleagues working on your project would not be good candidates.

The nature of the participants required for your study often play a role in this decision. Studies that involve systems for general use by a broad range of users should be able to attract a suitably large pool of participants, even if hundreds of people are needed. On the other hand, research aimed at studying very specific populations may need to rely on substantially smaller pools of participants: there simply aren't tens of thousands of potential participants for the study of a tool for space-shuttle astronauts. Studies of domain experts often face challenges in this regard.

Finding a suitably large participant pool can be particularly challenging for research involving people with disabilities. In addition to being an often-overlooked segment of society, people with disabilities often face significant challenges in transportation, making trips to research labs difficult. Studies with these users are often smaller, tending towards observational case studies with two or three users (Steriadis and Constantinou, 2003), rather than controlled experiments, see Chapter 15 for more details.

The time required for each participant is another important factor. Studies that require a single session of limited length (perhaps a few hours) can enroll larger numbers of participants than ethnographic observations that may involve several days or controlled experiments that require multiple sessions conducted over a period of weeks. As the time required from each participant – both in terms of direct involvement and the elapsed interval from start to finish – increases, it becomes more difficult to recruit and retain people who are willing to commit to that level of involvement.

How many participants should your study have? You should start by using your design as a guide. Ethnographies and case studies can be successfully completed with as few as two or three people. Numbers vary wildly for controlled experiments: although studies with as few as 12 users are not uncommon in HCI, results with 20 or more users are more convincing. From that base, you might expand to involve as many subjects as you can reasonably afford to include. You should then add a few more for pilot tests, replacements for participants who drop out, and a margin for error. Investigation of related work in the research literature can help in this regard: basing your population on a population used in similar prior work can be a good strategy. If there is no clearly related work, you might be able to use a smaller population.

14.1.3 Recruiting participants

Once you have determined *who* your participants are and *how many* you need, you must find them and convince them to participate.

If you work for a large corporation that frequently performs user studies, you may be able to draw upon the expertise of a dedicated group that maintains rosters of people interested in user studies and generates participant pools for research. Those who don't have such resources available (i.e., most of the professionals who conduct HCI studies) generally must do their own legwork.

The characteristics of your desired participants play an important role in determining how you will go about finding them. If you have relatively few constraints, recruiting is relatively simple. Advertisements and flyers on your college, university or corporate bulletin boards (both physical and electronic) can entice users. However, this must be done carefully: if you wish to get participants with a wide range of ages and education by recruiting on a university campus, you should be careful to explicitly recruit faculty and staff, as well as students. Notices in local newspapers and on community-oriented websites can be useful for recruiting an even broader group of participants.

More specific requirements are likely to require more focused recruiting efforts. Increased specificity in advertisements is a starting point: you might specifically indicate that you are looking for female college students. Community groups, professional organizations, and similar groups can be helpful for finding people with other, more specific characteristics. Many of these groups will be willing to pass messages along to members, particularly if the research may be of interest to them. If you can find a group of people that meet your specific

needs, it may help to go to them. If you can give a short presentation at a meeting and make yourself available for questions afterwards, you may encourage otherwise reluctant people to participate. Email lists and online groups can be helpful in this regard as well, but these tools should be used carefully: sending out messages that don't comply with posted group or lists policies is inappropriate. Sending unsolicited email messages directly to individuals is almost certainly a bad idea. Although an email message that comes from a trusted mailing list might be well-received, the same message sent directly by an individual might be seen as annoying junk email.

Focused ethnography and long-term case studies require fewer subjects, but the effort involved in enrolling each participant may be greater. These projects may require building cooperative arrangements with companies, schools, other organizations, and individuals in order to identify appropriate subjects. Many academic researchers address these challenges by bringing in outside organizations as collaborators. In addition to creating a formal agreement, collaboration can also provide funds that support the efforts of the cooperating organizations.

Incentives can often motivate people to participate. Many undergraduates have been lured into research sessions by promises of cash or pizza. If you can pay your subjects for their time, do so. Gifts can be more appropriate for some participants – particularly children. If you don't have enough funds to pay all participants, you can offer to enter them in a raffle for an MP3 player or similar desirable prize. Compensation can also be a motivator that can elicit desired behavior: in one study on interruption, researchers asked participants to both complete a memory task and respond to interrupting signals. In order to entice participants to complete both tasks, extra payment was given to the subjects with the best performance (Gluck, Bunt and McGrenere, 2007). Incentives for organizations that assist in recruiting can also be useful. In addition to the research collaborations described above, you might pay groups as consultants (see the Menu Task Performance Studies with Blind Users sidebar for an example).



Menu Task Performance Studies with Blind Users

Task performance with hierarchical menus has been the subject of many studies over the years, leading to a general consensus that menus with many choices at each of a few levels (broad, shallow trees) lead to faster task completion than menus with a few choices at each of many levels (narrow, deep structures), see Chapter 1. As these studies have generally been conducted with sighted users, who could rely upon a visual scan to quickly identify items in a long list, we were interested if these results would hold for blind users who rely upon the serial presentation of items by screen readers. To address this question, we designed a study based on an early experiment that looked at breadth vs. depth in web-based choices from an encyclopedia (Larson and Czerwinski, 1998).

Experimental studies involving blind people can be particularly challenging to run. As blind people often face challenges in transportation, expecting them to come to us would have been unrealistic. We also knew that we wanted a particular population: experienced users of a particular screen-reader package, who did not have any residual vision.

We enlisted the help of the National Federation of the Blind (NFB), who helped identify potential participants and provided us with access to space in their offices, where we were able to run the study. NFB was paid as a consultant on the project and study participants were compensated as well. Due to the specific nature of the participants, compensation was significantly higher than is customary for similar studies.

Compensation should be commensurate with the amount of time requested and the type of participant involved. Busy professionals may command a higher fee than students or children. For longer ethnographic or case studies, particularly with domain experts, direct payment for study participation is unlikely to account for the value of their time. In these cases, collaboratively funded research may be the best approach. For formative studies aimed at capturing requirements for systems to be used by domain experts, the ability to use the software being developed in their daily work might be a powerful enticement.

Special populations may require creative incentives and accommodations. If you are working with children, you might give them small toys as gifts for participating (cash compensation for accompanying parents is probably always welcome). Elderly people or others without easy access to transportation may be interested in participating but may be unable to make the trip to your lab or office. You might consider trying to conduct your study in participants' homes, community centers, or other locations that would be easy for interested participants to travel to.

Some studies may have additional requirements that require screening of interested participants to determine whether or not they meet important criteria. For example, tools designed for novices should probably not be evaluated by people who work professionally with similar interfaces. Initial questions and interviews with potential subjects can be important tools for ensuring that an individual is appropriate for your study. Specific questions about education, age, experience, and other important attributes can be asked to verify that there is indeed a good match. If you take this approach, you might also consider asking whether they are willing to be contacted in the future for subsequent studies. People who agree to future contact can form the basis for a home-grown database of study participants. Maintaining such a database may involve a fair amount of work, but it can be potentially very useful if you plan to run many studies.

Your database of potential subjects can be an important safety net in the event of difficulties along the way. You may start out with 15 (or 20, 30, or 60) participants with

confirmed appointments, only to find that several cancel at the last minute or simply fail to show up. Other problems associated with participant characteristics may force you to dig deeper for a wider range of ages, skills, or backgrounds. If the participants in your study of a general-purpose tool for managing personal photos are all men between 35 and 40 years old (or women over 60), you might have a hard time arguing that your results are indeed general. It's easy to argue that better planning and participant screening might help with this problem, but such details are often not obvious from the beginning. If you're faced with this dilemma, your best option might be to dig deeper into your list, inviting more participants to form a larger (and hopefully more representative) study.

Experiments that involve multiple experimental conditions may require dividing participants into roughly equal-sized groups. If you are comparing performance across user attributes – such as age or gender – your groups must differ in the relevant attributes, while remaining as comparable as possible for other characteristics. If your potential pool of participants is large, you need to select participants in a manner that minimizes any potential bias in selection: selecting the first names from a list that is sorted by gender may get you a group of subjects that is entirely male or female. See Chapter 4 for more discussion of these and related issues in population sampling.

14.2 Care and handling of research participants

Studies with human participants put researchers in a privileged position. As “scientific experts”, researchers have expertise, experience, and contextual knowledge that make them well-equipped to understand the reasons for conducting the experiment and the potential costs and benefits involved in participation in a study. Potential participants may lack some or all of this relevant background.

Research studies should be designed to protect participants. Informed consent – the notion that research participants should be provided with the information needed to make a meaningful decision as to whether or not they will participate – is the cornerstone of this protection. Academic and industrial organizations that conduct human subjects research generally rely on institutional review boards to review proposed research for any possible risks and to guarantee that appropriate procedures for informed consent are being followed.

14.2.1 Protecting participants

Participation in a research study involves multiple agreements between the participant and the researcher. The participant agrees to perform certain tasks as needed by the experiment and the experimenter frequently agrees to provide some incentive or compensation to the participant. Perhaps more importantly, experimenters agree to conduct responsible research that protects participants' rights, health, and safety.

Risks to participants are often most pronounced in medical research, where investigation of new drugs, devices, and procedures can lead to health risks, particularly when things don't work as intended (or hoped). However, physical harm is not necessarily the only relevant concern. Famous psychology experiments have shown how research that places

people in uncomfortable situations can cause significant emotional distress (see the Milgram's Experiment and Stanford Prison Experiment sidebars). Although some HCI experiments might raise these concerns, most of the studies in our field are low risk. Some studies may lead to fatigue (from mouse movements) or eye strain, but these risks are minor. Regardless of the level of risk involved, researchers must treat human subjects appropriately.

Research
in
Practice

Milgram's Experiment

Perhaps the most famous example of deception in psychology research, Stanley Milgram's obedience experiment illustrates one possible extreme of human subjects research.

In this study, subjects were told that they were participating in a study of the effect of punishment on learning. They were asked to administer tests to another subject – a “learner” – who would have to identify a word that had previously been associated with a stimulus word. Subjects were told that they had to administer an electric shock to the learner if incorrect answers were given and that the voltage of the shock should be increased after each incorrect answer. Shocks were described as being “extremely painful”, but incapable of causing permanent damage (Milgram, 1963).

This description was an elaborate deception aimed at concealing the true goal of the experiment: a study of the limits of obedience. As the “learner” was in fact a colleague of the experimenter's, no actual shocks were administered. However, the subject did receive a mild shock to provide evidence of the authenticity of the equipment and the learner acted as if shocks had been applied. The experimenter participated actively in the deception, urging subjects to continue with the experiment even when they expressed reluctance.

The results of the study were intriguing: of 40 participants, all continued giving shocks until after the point where the “learner” kicked on the wall and stopped responding to the test questions. Most (26 out of 40) of the participants administered the maximum level of shock – two steps beyond “Danger: Severe Shock.” Participation caused discomfort including nervous laughter, embarrassment, and seizures for several subjects.

This experiment would not have worked without deception: had the subjects known that they were not actually administering potentially painful shocks, they presumably would have been even less reluctant to participate. The deception created a scenario in which obedience had a real cost, in terms of the distress associated with inflicting harm on a fellow human being.

Milgram's experiment would probably not be considered appropriate human subjects research in most current research environments. The extreme nature of the psychological distress involved in these experiments and the strong reactions experienced by some of the participants raise serious questions as to whether such research can be conducted responsibly (Milgram, 1963).

Virtual environments provide interesting possibilities for subsequent investigations of similar phenomena without raising the ethical concerns associated with Milgram's experiment as originally executed. In a "virtual reprise" of those experiments, subjects were asked to administer shocks to a female virtual human in an immersive environment. The use of a computer-generated character eliminated the need for deceit, thus removing some of the possible ethical objections. Although participants knew that they were interacting with a computer-generated avatar, they responded to the situation as if they were working with a real person, particularly if they could see the avatar (as opposed to communicating via a text chat interface) (Slater *et al.*, 2006).

Research
In
Practice

The Stanford Prison Experiment

Many interesting and important questions about human behavior in difficult situations can only be examined by conducting studies that expose participants to the risk of significant psychological distress. As interesting as these questions may be, the risks are substantial enough to make this research effectively off limits.

The Stanford prison experiment, conducted by Philip Zimbardo and his colleagues during the summer of 1971, provides an example of both the risks and insight potentially associated with research that exposes participants to significant emotional distress. In order to examine the social forces associated with prisons, the researchers divided a group of Stanford undergraduates (all males) into "guards" and "prisoners". Prisoners were arrested at their homes, blindfolded, placed in uniforms, and incarcerated in a makeshift prison constructed in the basement of Stanford's psychology building. Guards were not given training – they were simply told to do what was necessary to maintain order.

The researchers and participants were all surprised by their responses. Both guards and prisoners completely fell into their roles. Guards humiliated prisoners, using tactics such as awaking prisoners throughout the night for "counts" and placing people in solitary confinement to establish their authority and prevent rebellion. Prisoners temporarily lost their personal identity, thinking of themselves only by their prisoner number. They were passive, depressed, and helpless. One prisoner suffered significant stress, including crying and rage. Both the guards and the researchers responded like real prison staff, believing that he was faking. Dr. Zimbardo – the professor in charge of the experiment – found himself acting like a prison warden, bristling at concerns for the well-being of the prisoners – who were, after all, innocent bystanders. Originally planned for two weeks, the study was terminated after six days, out of concern for the participants (Haney, Banks and Zimbardo, 1973; Zimbardo, 2008).

The observation that seemingly ordinary people would quickly assume the role of sadistic prison guards raises serious questions about the role of context in determining human behavior. Although we would all like to think that we would not behave abusively in such contexts, the Stanford Prison Experiment raises the concern that environment and expectations can play a huge role in encouraging seemingly inhuman behavior. This lesson continues to have significant relevance: Philip Zimbardo has been an oft-quoted commentator on the behavior of guards at the Abu Ghraib prison in Iraq (Zimbardo, 2008).

The Stanford prison experiment also provides a cautionary tale regarding the evolution of research ethics. Despite the known potential for harm, this study was approved by Stanford's Human Subjects Review Board, participants signed an informed consent form, and a 1973 review from the American Psychological Association determined that the study had been consistent with existing ethical guidelines (Zimbardo, 2008). Changing views on responsible research – influenced at least in part by this – have led to a much more conservative view of appropriate research. Philip Zimbardo publicly apologized for his role in the study (Zimbardo, 2008) and the establishment of beneficence – maximizing of benefits while minimizing harm (National Commission, 1979) – argued for research that would strive to avoid the harms seen in the prison experiment. It's hard to imagine a study with this degree of potential harm being approved by any modern institutional review board.

Specific definitions of the responsibilities of researchers grew out of concerns about inappropriate medical procedures conducted during the mid-20th century (see the Informed Consent: Origins and Controversies sidebar). In 1979, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research published the Belmont Report (National Commission, 1979). This document established three principles for the treatment of research participants: respect for persons, beneficence, and justice. Respect for persons involves allowing individuals to make independent and autonomous decisions regarding their participation in research. Researchers must allow participants to make judgments and must provide the information necessary for making those judgments. Special consideration must be given in cases of illness or disability that may limit an individual's ability to make independent decisions. Beneficence refers to the need to minimize possible harm while maximizing possible benefits. Justice requires that neither the burdens of participating in research nor the benefits of the research should be limited to certain populations, particularly when some groups of people may be easily manipulated (National Commission, 1979). These principles form the basis for informed consent.

Informed Consent: Origins and Controversies

Famous (or infamous) medical research experiments conducted during the mid-20th century led to the development of modern concepts of informed consent and appropriate treatment of research participants. Nazi Germany's use of concentration camp prisoners in often brutal and barbaric medical experiments led to the Nuremberg code, which established some of the principles behind informed consent (National Cancer Institute, 2001).

The US Public Health Service Syphilis Study at Tuskegee involved hundreds of black men with syphilis over 40 years. Although they were told that they were being treated, no treatment was in fact given, and efforts were actively made to prevent participants from getting treatment (Centers for Disease Control and Prevention, 2007). Several other studies in the US involving administration of drugs or treatment without consent were conducted in the US after the end of World War II (Pellegrino, 1997). More recently, drug trials conducted by Western companies in countries such as India have raised concerns about the nature of informed consent across such cultural and financial divides (Sharma, 2005).

The costs associated with these studies are not limited to the substantial harm inflicted upon the subjects. These unethical experiments reflect poorly on science and scientists in general, harming public trust and increasing reluctance to participate. One study of both white and black residents of Detroit found that black residents were more likely to have heard of the Tuskegee experiments. They were also more likely to be distrustful of researchers and less likely to participate in research (Jones, 1993; Shavers, Lynch and Burmeister, 2000).

Participants should also be assured that their privacy will be protected. Work in the field of privacy protection provides guidance that can help HCI researchers protect the privacy of study participants (Patrick, 2007b). Researchers should obtain consent for the collection and storage of personal information; limit the information collected to that which is necessary; identify the uses that will be made of any information; limit the use, disclosure, and retention of the information; securely protect any information; disclose policies and procedures; provide a means for addressing concerns regarding compliance with information practices; and be accountable for those practices (Patrick, 2007b). Patrick (2007b) provides questions that can be asked in each of these areas to guide privacy practices.

The use of photography and video or audio recording presents special challenges regarding the privacy of participants. Photos, videos, and audio recordings can be very useful tools for illustrating the use of an interface, but they can also unambiguously identify individuals as having participated in a research project. There are several steps that you should take in

any project before you start the shutters snapping or cameras rolling. You should clearly tell participants what you are recording and why. If you are going to consider using images of participants in any publications or reports, participants should be fully informed of this possibility. These practices should be mentioned in your informed consent forms (Section 14.2.2) and discussed with participants. If you are video-recording, you might consider recording a portion of the discussion, taking care to include footage of the participants explicitly agreeing to be video-recorded. You should plan your photos or videos carefully: if you are really interested in what is going on with the interface, take pictures and video of the inputs and display – not the faces of the participants. You might be able to shoot over the users' shoulders to get a fuller view without identifying your participants. Similarly, audio recordings captured for potential distribution should minimize use of the participant's voices – record the voices of the research staff if necessary. If you must show people in action, you might consider using image-manipulation techniques, such as blurring or black bars over the eyes to hide the identity of the participants. Pictures or videos of the research staff might be more appropriate for distribution. Finally, you should provide an alternative for participants who are concerned about their privacy: you probably don't need video or audio recordings of every individual in your study.

14.2.2 Informed consent

The notion of informed consent has two parts. "Informed" means that study participants must understand the reason for conducting the study, the procedures that are involved, potential risks, and how they can get more information about the study. Without this information, participants do not have the information necessary to make a truly meaningful decision as to whether or not they wish to participate. If potential participants are not told that the use of a specific virtual-reality environment can occasionally cause nausea, particularly sensitive individuals may agree to participate without being aware that they might be subjecting themselves to an unpleasant experience. For these reasons, researchers should strive to clearly provide information that is relevant and necessary for appropriate decision-making. Truly informing potential participants means that the information must be provided in a manner that is comprehensible (National Cancer Institute, 2001). The reason for the study, the procedures being used, and other details should be provided in a manner that is clear, accessible, and free from professional jargon.

The second, equally important notion is "consent": participation in research studies should be entirely voluntary and free from any implied or implicit coercion. Potential participants should not be given any reason to believe that a decision not to participate will lead to repercussions or retaliation, whether in the form of punishment by employers; withholding of medication or the use of a system; or disapproval from the researcher. Researchers in academic settings should be very careful about giving students credit for coursework in exchange for their participation in studies: if an alternative means of earning

the credit are not provided, some students may feel that their grades will suffer if they decline to participate. In such circumstances, participation would be coerced, not consensual.

In most cases, researchers provide participants with an informed consent document that contains several sections (National Cancer Institute, 2001).

- *Title and Purpose:* Why is the study being done?
- *Description of Procedures:* What will be asked of participants? For HCI studies, this probably involves using one or more interface variants, discussing goals and needs, commenting on design proposals, and other related tasks.
- *Duration:* How long will each participant be involved in the study? This should tell the user how much time will be involved. If there are multiple sessions, the number of sessions, the length of each session, and the elapsed interval required should all be specified.
- *Risks:* What risks might be involved in participation? Medical trials may involve the risks of unknown drug side-effects, but the risks are generally less severe in HCI studies. Fatigue, boredom, and perhaps slight discomfort due to repetitive motion are possible risks for studies involving desktop computers. Virtual-reality systems may involve some risk of nausea or disorientation. Studies involving mobile devices, computers in cars, or other interfaces in non-traditional settings may involve additional health or safety risks. Evaluation of the potential distractions caused by computing devices in cars should probably not be conducted in cars driving on public roads! Other interfaces involving social interactions may pose emotional risks, if tasks or content may prove upsetting to participants (see Milgram's Experiment sidebar). The privacy risks of photography and video or audio recording are discussed above; projects involving online-conferencing or ongoing use of online chat systems may present similar concerns. Experimenters should, of course, design studies to minimize all risks. Any remaining risks should be described in detail in informed consent forms and then discussed honestly and thoroughly with study participants.
- *Benefits:* What are the benefits of participation? Some researchers may provide participants with ongoing access to software that is being evaluated. In other cases, financial or material compensation is the main benefit.
- *Alternatives to Participation:* What other options are available? For most HCI studies common alternatives include simply not participating and continuing to use the software that was being used before the study.
- *Confidentiality:* Participants' privacy should be respected. This section of the form generally includes comments indicating that personally identifying information will not be used or published in any way. Confidentiality is a particularly important issue for HCI research involving observation of user behavior such as search or information use activity. Web search, email organization, and other activities may reveal sensitive personal information that could compromise confidentiality. Proper protection of participant privacy involves limiting the use, disclosure, and retention of data; taking appropriate measures

- to protect data, including encryption and secure storage; openly describing policies and practices; providing avenues for challenging compliance with data protection procedures; and providing for training and related measures to ensure accountability (Patrick, 2007b).
- *Costs/Additional Expenses:* Are there any financial expenses or other costs associated with participation? Although such costs may not be inappropriate, they may discourage some users from participating. If you are going to ask participants to make costly trips to travel to your location, to purchase software for their computer, or to spend significant amounts of time entering data into diaries, you need to make sure that they are aware of these costs.
 - *Participant's Rights:* This section should make three important points:
 - Participation is voluntary.
 - Participants can choose to stop participating at any time, without penalty.
 - Participants have the right to be informed of any new information that will affect their participation in the study (National Cancer Institute, 2001).
 - *Contact Information:* Who should participants contact if they have questions or concerns? This section should contain names and contact information for the researchers in charge of the study, as well as for representatives of the institutional review board or other appropriate body.
 - *Supplemental Information:* Where should participants go for further information? This section should list resources that can be used for additional information, including (but not limited to) descriptions of the research program and institutional policies and procedures for research involving human subjects.
 - *Signature:* Participants should sign a copy of the consent form. The signature should be accompanied by a statement indicating that the participant:
 - has volunteered to participate;
 - has been informed about the tasks and procedures;
 - has had a chance to ask questions and had questions answered;
 - is aware that he/she can withdraw at any time;
 - consented prior to participation in the study (Shneiderman and Plaisant, 2009).

The researcher should provide a copy of the consent form to each participant for reference, while retaining the signed copies as documentation of the consent.

Construction of an informed consent document can be a useful step in ensuring that your research meets accepted ethical standards. If you have accounted for the risks, benefits, alternatives, and confidentiality measures associated with your project, the relevant sections of the document should be relatively straightforward to put together. Similarly, difficulty in construction of these sections may indicate the need to rethink proposed practices in procedures.

Writing clear, concise informed documents is not trivial. One study of informed consent forms for medical research studies found that users preferred simpler statements written at a seventh-grade level (as opposed to at a college graduate level) but the simpler statements did not lead to greater comprehension (Davis *et al.*, 1998). Pilot testing of the consent forms,

either as part of a pilot test for an experiment or via reviews by potential participants or collaborators can help identify confusing language or areas that may need clarification. An example informed consent form is given in Figure 14.1.

Informed consent requires affirmative agreement from an individual who is capable of understanding the implications of agreeing to participation in the research. Research involving participants who are not able to interpret informed consent forms may require additional measures. When children participate in research studies, parents or legal guardians are generally asked to consent to the participation. When possible, children may also be asked to “assent” – to agree to participate – even if they are not capable of giving informed consent (Society for Research in Child Development, 1991). This assent would be in addition to – not instead of – parental consent. Considerations of informed consent and users with disabilities are discussed in Chapter 15.

Local or national legislation may place additional constraints on the content of an informed consent document. In the United States, federal regulations prohibit language in informed consent forms that would waive legal rights or absolve researchers of legal responsibility (National Cancer Institute, 2001).

The use of informed consent forms – even those that are approved by institutional review boards (see Section 14.2.3) should not be seen as a green light to move forward with research that may otherwise raise questions regarding respect for the rights and concerns of participants.

14.2.3 Institutional review boards

Universities, hospitals, corporations, and other organizations that conduct research often have standing committees that review and approve projects involving human subjects. These institutional review boards (IRBs) examine proposed studies for appropriate practices, procedures, goals, and disclosures. By conducting this review prior to the start of human subjects research, IRBs protect all of the groups and individuals that may be affected by the research. Participants are protected by examination of proposed research for any elements that may be manipulative, coercive, or otherwise abusive. Proposals that contain any such elements should not be approved by IRBs. Researchers and institutions benefit from the knowledge that the proposed research has been reviewed for issues that may cause embarrassment or legal liability. Although this review is certainly not foolproof, it generally works well in practice.

IRB review and approval for proposed research generally begins when a researcher submits materials relating to proposed research. A description of the proposed research, draft informed consent forms, instructions to be provided to users, questionnaires, and materials to be used during the course of the research are some of the items that might be required. Upon receipt of these materials, the IRB will review them for completeness and content. The board may approve the research, request additional information, require revision of materials, or take other steps as appropriate.

INFORMED CONSENT FORM

Evaluating Menu Selection Task Performance

PRINCIPAL INVESTIGATOR: A. Researcher
 Department of Computer and Information Sciences
 Research University
 Phone: 555-555-5555
 Email: researcher@research.edu

Purpose of the Study: The goal of this study is to understand how computer interfaces might be customized to best suit the needs of users. Participants will be asked to use a menu interface to find items in various multi-level hierarchy designs. Task completion times and subjective responses will be used to determine which (if any) design is most suitable for these users.

Procedures: Participation in this study will involve two phases. In the first phase, you will be asked to use a web browser to make selections from a menu of choices, in order to locate a specified entry. You will be given the opportunity to try a sample task, and then you will have to complete multiple tasks with different menu structures. This study should take about one hour to complete.

After you have completed the experimental tasks, we may ask you some questions about the various interfaces. These questions will be designed to help us understand which (if any) of the interfaces you preferred, and why. We may also ask some general questions about your habits and practices with respect to computer use.

Risks/Discomfort: You may become fatigued during the course of your participation in the study. You will be given several opportunities to rest, and additional breaks are also possible. There are no other risks associated with participation in the study. Should completion of either the task or the interview become distressing to you, it will be terminated immediately.

Benefits: It is hoped that the results of this study will be useful for the development of guidelines for the design of user interfaces that will help people use computers more effectively.

Alternatives to Participation: Participation in this study is voluntary. You are free to withdraw or discontinue participation at any time.

Cost and Compensation: Participation in this study will involve no cost to you. You will be paid for your participation.

Confidentiality: All information collected during the study period will be kept strictly confidential. You will be identified through identification numbers. No publications or reports from this project will include identifying information on any participant. If you agree to join this study, please sign your name below.

_____ I have read and understood the information on this form.

_____ I have had the information on this form explained to me.

 Subject's Signature

 Date

 Witness to Consent Procedures

 Date

 Principal Investigator

 Date

If you have any questions regarding this study please contact Dr Researcher at (555) 555-5555 or the Institutional Review Board Chairperson, Dr Chair Person, Research University, (555) 555-6666.

Figure 14.1 Informed consent form.

As research cannot begin until the IRB approval is complete, it is generally best to start this process early. Some research funding agencies will not release any funds until appropriate IRB approvals have been obtained. As each IRB has its own rules, it is important that researchers understand and follow the appropriate procedures for their institution. Many IRBs have websites that describe policies and provide relevant forms. It's a good idea to familiarize yourself with this material. Although some boards consider applications on a rolling basis, others have scheduled meetings, with published submission deadlines for consideration at each meeting. Attention to detail is particularly important for boards that meet on a set schedule: if your IRB meets bi-monthly, minor omissions in a proposed package may lead to a two-month delay in acquiring the necessary approval.

Some IRBs – particularly those at large research institutions with affiliated medical schools – may spend much of their time focusing on drug or treatment studies. If your IRB falls into this category, board members may not be aware of the techniques used in HCI research (as described in this book). You may have to spend some time and effort explaining ethnography, research based on online data sources, or other techniques that they are not familiar with. If you run into this sort of challenge, you should stress the widespread application of these techniques, and the existing body of research from groups such as the Association of Internet Researchers (www.aoir.org) or the Ethnographic Praxis in Industry Conference (EPIC). It's best to approach such discussions from a collegial, not confrontational, perspective.

Although the paperwork required by some IRBs may feel like a nuisance, you should consider your IRB as an ally. By insisting upon procedures, IRBs protect researchers and institutions from problems associated with research that goes wrong. IRBs can also provide helpful feedback in situations that may raise questions. Some projects may blur the lines between participating in the research and acting as a collaborative partner. For example, projects involving participatory design may involve ethnographic observation of users in the workplace. Is informed consent necessary in this case? Although the conservative approach of requiring informed consent is unlikely to be inappropriate, discussing this question with a member of your IRB might provide insight into your institution's policies regarding such research. Many IRBs require researchers to take training courses before conducting any studies involving human subjects research. These courses may not seem exciting, but they can provide valuable information that might prove helpful when you are preparing informed consent materials.

Organizations that infrequently engage in human subjects research may not have an established institutional review board. This may be particularly true for small companies that run occasional user studies. If you find yourself in such a situation, it may be helpful to discuss matters with appropriate professionals in your organization, including community relations staff and legal counsel. IRBs from nearby research institutions may be willing to provide feedback as well. The use of informed consent forms and proper procedures is always appropriate, even in the absence of a formal review from an IRB.

14.2.4 Potentially deceptive research?

Researchers may occasionally have legitimate reasons to be less than forthcoming about the goals and procedures of their research. The practice of potentially deceptive research involves asking a user to perform a set of tasks that are described as relating to a particular goal, when the researcher is actually interested in addressing a different question unrelated to the goal presented to the user. Although concealing the true nature of the study does present some concerns regarding the validity of informed consent, this practice is often necessary, particularly in situations where full disclosure might compromise the realism of the study.

A study involving security and usability provides an example of the use of deception in HCI research (Schechter *et al.*, 2007). This study had two goals: to determine the influence of security feedback and to see if participants using their own data would behave more or less securely than those who were role-playing using someone else's data. As the researchers were concerned that study participants would not behave naturally if they were told that usability was being studied, they were told that the purpose of the study was to "help make online banking better" (Schechter *et al.*, 2007). Participants were asked to perform online banking tasks. Some participants were "role-playing" – they were asked to pretend that they were a specific individual with specific goals in mind; others used their own bank accounts. In addition to finding that security indicators were not particularly helpful, this study found that people using their own data behaved more securely than those who were role-playing (Schechter *et al.*, 2007).

Schechter *et al.* (2007) used deceit as a means of setting up conditions that maximized the realism of the experiment. By presenting users with real online banking tasks, they focused the experiment on how actual users might behave when using online banking on their own. If participants had been told that the experiment was examining their behavior regarding security and privacy, they might have paid extra attention to their behavior in these areas. This use of deception may be useful and valid, but it does have its limits. These limits arise from the established psychological concept of demand characteristics (Orne, 1962), which states that participants in a research study may act in a manner that attempts to validate the hypotheses being tested. In this study, participants may have taken the goal of improving online banking to heart, perhaps acting more insecurely than they otherwise might have (Patrick, 2007a).

Deception in HCI research should be used carefully and sparingly. As deception pushes at the limits of the concept of informed consent, researchers should be careful to frame deceptions clearly, justify their use, and minimize any risks – particularly regarding discomfort and distress – that may be involved (See the Milgram's Experiment sidebar for a famous example of deceptive research). Participants in studies involving deception should be thoroughly debriefed at the end of their participation. Debriefing has been shown to help deceived participants eliminate negative effects and even to have experiences that were more positive than those of participants who have not been deceived (Smith and Richardson, 1983).

14.2.5 General concerns

Participants are crucial to our studies – without them, HCI research would be all but impossible. We should make every effort to treat participants in a manner that reflects this importance. Compensation for time and effort is certainly helpful, but researchers should also take concrete steps to make participation convenient and enjoyable. Comfortable surroundings may put participants at ease. Ample opportunities for rest or bathroom breaks should be provided, particularly for studies that involve longer research sessions. Flexibility in scheduling and location can be particularly important for some users: enrolling professionals in your study may require that you travel to their workplace or allow for sessions outside of traditional working hours. If your study is fun and convenient, participants may be more likely to help your recruiting efforts by urging friends and colleagues to join in.

These concerns are particularly important for special cases that place a significant burden on participants. Longitudinal studies require participants to make a huge time commitment – many hours over weeks or months. Research on people with disabilities may require enrolling participants who have significant difficulty traveling. You may find that engaging the required range of participants requires traveling to participants' home or workplace, at times of their choosing.

When working with human participants in any form of HCI research, you must pay careful attention to your role as a researcher. Participants may be impressed or intimidated by your presence, your use of language, your technical skills, the context of the experiment, or any of a variety of related factors. This is particularly true for observations and contextual-inquiry, where you will spend a great deal of time in close contact with one or more participants. Although you should make every reasonable effort to help participants feel as at ease as possible, you should also be aware that your presence may have an impact on observed performance. In some cases, participants may exhibit the “demand characteristics” described above, trying to behave in the manner that they think you are looking for.

Others have claimed that the mere act of participating in an experiment will influence user behavior, in the so-called “Hawthorne effect”. Although this effect has been the subject of significant debate among scientists, some suggested responses are clear and appropriate. Researchers should never give feedback regarding user performance during the course of a study and experiments involving the comparison of multiple interfaces should be controlled and “blind” – participants should not know if one of the alternatives is favored by the researchers (Macefield, 2007).

More generally, these concerns about the influence of researchers on experimental results point towards a need to be modest about the results of our research. All experiments have flaws and no single study establishes incontrovertible facts on its own. When reporting results and drawing conclusions, we should avoid overstatement, admit the flaws in our research, and point the way for future work that will bring greater understanding.

14.3 Online research

Working with human subjects is often challenging. Scheduling sessions, recruiting participants, finding appropriate space, and managing other logistical details require time and energy, neither of which ever seems to be available in abundance.

Online HCI research presents the tantalizing prospect of a way out of these challenges but, as you might have guessed, there is no silver bullet. This section outlines research issues including appropriate topics for online research, recruiting, study design, ethical concerns, and methods for data collection.

14.3.1 Appropriate topics for online research

Although it may seem somewhat obvious to note that online research will involve working with participants who are online, this helps point us toward the insight that online HCI research may be most appropriate for studies about the tools that people use online and the uses that they make of those tools. Participants in online studies will probably be working with web browsers, chat tools, and related online software as they read instructions, provide informed consent, perform tasks, and otherwise complete your experimental protocol. Research that works within this realm may be most successful.

As far as tools are concerned, this implies that studies involving web applications or online tools may be particularly well suited for online research. If you are interested in testing the usability of website design or using a dynamic website to collect data on task performance, an online study may be very appropriate. If you are running the website on your own servers, web logs (Chapter 12) can provide useful feedback regarding timing, tasks, and errors. Conversely, studies of other application software, mobile devices, or novel interaction devices may be harder to do online: data collection is likely to be more difficult, incompatibilities between software versions may pop up, etc.

That's not to say that online studies of website designs are easy. Good design practice certainly calls for cross-platform testing, but there is no guarantee that you won't run into versioning and compatibility problems, even with seemingly straightforward web pages. If your test involves dynamic Javascript and HTML combinations, bugs and plug-ins could cause all sorts of trouble.

Investigation of the uses that people make of online tools might involve ethnographic analysis of online bulletin boards for various communities of interest (Maloney-Krichmar and Preece, 2005) and other studies that attempt to understand online socialization, resource usage, and other behavior. The Association of Internet Researchers (www.aoir.org) hosts an annual conference with numerous studies along these lines, many of which are of direct interest to the HCI community.

14.3.2 Recruiting

By opening your research up to the Internet, you provide yourself with access to a much larger pool of participants. Recruiting can be easier, as emails to appropriate lists and postings on

various websites can go a long way towards identifying potential subjects. As online research generally involves the use of a website or other online software, participants do not need to be local. Self-driven website or study tools allow participants to complete tasks at their leisure, eliminating the need for scheduling.

Just as the use of undergraduates as study participants introduces a bias that may not be appropriate for some studies, online recruitment limits your subject pool to a particular segment of the larger population: Internet users who are interested enough to participate. This may mean that you do not attract relatively inexperienced individuals or participants who limit their time online to relatively focused activities. Whether or not this poses a problem depends on the specifics of the study in question.

In some cases, online research can give you access to pools of participants that otherwise would have been unavailable. This is particularly true for people with disabilities, who may find traveling to a researcher lab to be logistically unfeasible (Petrie *et al.*, 2006), and domain experts, who may be hard to find in sufficient numbers in some locales (Brush, Ames and Davis, 2004). See Chapter 15 for more details on HCI research involving people with disabilities. Collaborative research involving distant partners can also be substantially aided by online tools for communicating and gathering data.

One important difference between online and in-person research is the potentially complete anonymity of participants in online studies. When you meet a participant face-to-face, you can usually make a pretty good guess about their age, gender, and other demographic characteristics. The lack of face-to-face contact with online participants makes verification of such details harder – you have no way of verifying that your participants are male or female, old or young. This presents some recruiting challenges, particularly if your research requires participants that meet certain demographic constraints such as age or gender. If your only contact is via email or other electronic means, you may not be able to verify that the person with whom you are communicating is who he or she is claiming to be. Online studies that don't require the participants to reveal their true identity (relying instead on email addresses or screen names) are highly vulnerable to deception. Certain incentives, such as offering to enter participants in a draw for a desirable prize, might compound this problem. For example, a survey aimed at a specific demographic group might draw multiple responses from one individual, who might use multiple email addresses to appear as if inquiries were coming from different people. Possible approaches for avoiding such problems include eliminating incentives; requiring proof of demographic status (age, gender, disability, etc.) for participation; and initial phone or in-person contact in order to provide some verification of identity. Since payment or other delivery of incentives often requires knowing a participant's name and address, verification of identity is often not an added burden.

Online research involves giving up a certain amount of control over both the participants and the process. When you meet participants face-to-face, you can gain a great deal of information by observing their actions and behavior. To varying extents, you can tell if they

are being truthful about demographic information, observe their subjective reactions to their participation, provide assistance when appropriate, and make note of any cues or observations that may seem pertinent. The contextual feedback associated with online research is much less limited. Even if you are doing synchronous research with video chat and screen capture, you will still be somewhat limited in the information that you will be able to observe during the course of the session.

14.3.3 Study design

Surveys (Lazar and Preece, 1999), usability evaluations (Brush, Ames and Davis, 2004; Petrie *et al.*, 2006), and ethnographic studies of support groups (Maloney-Krichmar and Preece, 2005) have all been successfully completed online. Recent examples of online usability studies have shown that both synchronous studies with domain experts (Brush, Ames and Davis, 2004) and asynchronous studies with disabled users (Petrie *et al.*, 2006) have yielded results comparable to those that were found in traditional usability studies. Perhaps due to difficulties in sampling and controls, online empirical studies of task performance are less common. One study of the influence of informal “sketch-like” interfaces on drawing behavior used an online study as a means of confirming the results of a smaller, traditional study. Results from the 221 subjects in the online study were highly consistent with the results from the 18 subjects in the traditional, controlled study in the lab. The agreement between the two sets of results provides a more convincing argument than the lab study on its own (Meyer and Bederson, 1998).

Opinions differ on the appropriateness of online research for different types of data collection. The lack of controls on the participant population might be seen as a difficulty for some controlled, empirical studies. Others have argued that as online research does not allow for detailed user observation, it is more appropriate for quantitative approaches (Petrie *et al.*, 2006). In the absence of any clear guidelines, it is certainly appropriate to design studies carefully and to clearly describe and document the reasoning behind any designs that are adopted. When possible, hybrid approaches involving both in-person and online research may provide additional data and avoid some of the downsides associated with each approach.

14.3.4 Ethical concerns

Although the usual guidelines regarding protection of participants apply to online research, numerous confounding factors can create some interesting and challenging dilemmas.

Studies of online communities must consider questions of privacy and online consent. What is the expectation of privacy when participants in an online forum post messages publicly? Are such messages fair game for researchers? Is informed consent required before messages can be used? What if the site is only accessible to users who register and login? These questions have generated debate, discussions, and some guidelines (Bruckman, 2002; Frankel and Siang, 1999), but specific issues vary from case to case. Creating communities

specifically for research purposes can be a successful – if not always practical – strategy (Bruckman, 2002).

Debriefing and informed consent online can also be tricky. Providing important information for either of these tasks via online text may not be sufficient. In-person studies provide the possibility of direct feedback: experimenters know if participants have any questions or if there is any post-experiment distress. These factors are much harder to gauge online (Azar, 2000). Although one study indicated that comprehension of informed consent forms online may be comparable to comprehension of forms on paper, poor recall in both cases illustrates the general challenge of constructing effective consent forms (Varnhagen *et al.*, 2005). These studies should not be undertaken without careful attention to IRB processes and approval.

Further complications in informed consent and debriefing arise with online studies involving deception. A series of studies of “phishing” – the use of forged emails to attempt to entice users to login to fraudulent websites, thus giving attackers access to their user names, passwords, and related credentials – used social network analysis and related means to identify potential participants, who were sent phishing emails. These emails effectively enrolled recipients in the study, without any prior knowledge or informed consent. Although the methods received IRB approval, these studies raised many concerns and controversies, including legal ramifications, potential for harm due to online debriefing, and technical issues relating to Internet hosting of study materials (Finn and Jakobsson, 2007).

The considerable challenges and headache associated with deceptive online research provide a strong argument against this sort of approach. If you find yourself tempted to try this sort of study, consider a lab-based study instead. You may still use deception in this case but the use of prior informed consent can help you avoid many difficult questions.

As with any HCI research, online research can be particularly challenging if there is potential harm involved or when dealing with special cases, such as research involving children. Technical measures such as encryption of transmitted data may be useful for privacy protection and for verifying parental consent in the case of minors (Kraut *et al.*, 2004). Laws such as the Children’s Online Protection Act in the United States may limit the amount of information that can be collected from minors. Researchers working in these areas should construct study materials carefully; consult with IRBs and external experts to review proposed procedures; and use traditional studies as opposed to online studies when appropriate (Kraut *et al.*, 2004).

14.3.5 Data collection

Web logs or other software designed to collect appropriate data and send it back to a server can be a powerful means of collecting experimental data in a manageable and accessible format. In addition to indicating which pages were visited and when, logs can provide information regarding the browser that was used and the “referring site” (where the user came from). Particularly when used with tools designed to extract and analyze patterns from such logs, this data can provide a useful picture of how websites are used. Commercial packages for remote usability evaluation provide similar functionality.

Similar techniques can be used with other software packages. In a process known as “instrumentation”, custom software tools can be extended to collect data and send it back to a remote server via a network connection. For example, a browser plug-in might be used to track mouse paths on a web page. This information might be sent to a server and correlated with web logs to provide an understanding of where a user’s mouse went on a given page. Other tools might be augmented with “talk-back” mechanisms, which might periodically send data back to a server – either silently or after alerting a user. Many of the automated data collection techniques discussed in Chapter 12 can be applied to online research.

Used appropriately – that is, with relevant disclosures to participants and safeguards to protect privacy – remote data collection tools can provide a wealth of data for online research.

Summary

Working with human subjects is one of the most challenging and informative aspects of HCI research. Finding appropriate participants; informing them of their rights; protecting their privacy; and answering their questions can be time-consuming and often tedious, but the results are more than worth the effort. Even when study participants criticize our designs or fail to confirm our cherished experimental hypotheses, they provide invaluable insight that provides a rigorous foundation for our work.

Whatever type of study you are running, it is never too early to plan for recruiting, informed consent documentation, and other aspects of human participation. Proper planning will keep your study from becoming one of the many that have been delayed by unforeseen circumstances including difficulty in finding participants, or delays in IRB approval.

Recruiting entails finding the right number of the right kinds of participants. For usability studies, ethnographic observations of users, interviews, focus groups, and other approaches aimed at gathering requirements or evaluating design proposals, this may mean understanding the audience of users and identifying a sample of participants that is broad enough to reflect the needs and behavior of potential users. *Designers* and *professional developers* conducting research of this sort might work with collaborators, marketing teams, professional organizations, or others with appropriate understanding and context to identify both the range of viewpoints that would be needed and possible sources of the appropriate individuals.

Empirical studies require consideration of both the diversity of potential participants and any confounding factors that might contribute to performance differences. Characteristics of desirable participants might both be informed by and influence experimental hypotheses. *Students* and *researchers* conducting these studies should be careful to plan their data analysis and recruiting together, to ensure that the participants will be selected to increase the power of the statistical analysis.

Appropriate respect for participants is a cornerstone of all research involving human subjects. Although *designers* and *developers* may not be required to secure the approval of institutional review boards, they should still endeavor to protect their subjects from any form of harm and to treat them with respect and dignity.

These issues are particularly relevant for studies that involve deception. Even when not required by institutional policy to do so, designers and developers would be advised to use formal informed consent forms to help participants make informed decisions. *Students and researchers* should take the time – again, as early in the process as possible – to understand the regulations in force in their institution, and to make sure that their approvals are in order before starting any project.

Designers and developers may find online studies to be an attractive means of evaluating proposed interface designs. *Students and researchers* will undoubtedly continue to find the prospect of online research too enticing to resist. Before moving studies online, HCI professionals should be careful to validate that their proposed designs will provide the desired information. Pilot tests may be particularly useful in these cases.

Human subjects research in HCI can be an unpredictable and often unsettling process. Unforeseen problems, including misinterpreted tasks and goals, systems failures and missed appointments, are routine: it's rare that a study (of any sort) goes off completely without a hitch. These matters can complicate data collection and interpretation: if a user chooses an interpretation of a written task that differs from your intent and then completes the task correctly, how do you interpret the result – is it correct or not? What should be done with results from a user who decides to withdraw from a study after completing only a portion of the tasks? As hard and fast rules for handling situations like these are few and far between, you may have to handle each issue on a case-by-case basis. The specific decisions that you make may be less important than how they are enforced: consistent application of policies and procedures will ensure your ability to make meaningful comparisons.

All participants in HCI research study should be well-treated and approached with an open mind. Participating in HCI studies should be fun and engaging whenever possible: by making our studies positive experiences, we encourage people both to participate and to provide useful feedback. As researchers, we should “expect the unexpected”: software will crash, devices won't work, and (perhaps most distressingly) users will hate our beloved inventions. High-quality HCI research takes these setbacks in its stride, all the while striving to observe carefully while maintaining respect for the people who give a bit of their time to help our studies along. By watching and listening carefully, we can learn from what they do and how they do it. That, after all, is the point of conducting user studies.

Discussion Questions

1. University researchers occasionally ask students in a class to participate in research studies. However, this practice may involve elements of coercion, as students may be concerned that refusal to participate may negatively impact their grade. Is voluntary informed consent possible in such a situation? What steps might be taken to reconcile the researcher's need for subjects with the students' right to decline to participate?

2. The virtual reprise of Milgram's experiment (see Section 14.2.1) asked participants to inflict harm upon a computer-generated avatar. This approach eliminates some of the potential ethical concerns associated with the original experiment, but may raise additional questions. As user behavior was similar to what was observed in the original experiments, it is possible that participants in the "virtual" versions would experience similar patterns of nervousness and distress. Do you consider this sort of research to be appropriate? What might be done to protect participants in this sort of experiment?
3. As part of a larger study of how various aspects of interaction in online worlds, such as Second Life, impact the offline lives of participants, you are interested in observing participants both online and offline. As you know, participants in online games such as these may not represent a broad cross-section of society. The race and sex of online characters may not reflect those of the real individuals involved and some may choose to hide their "real" identity. Given these challenges, how might you go about finding a group of participants that would be interesting to work with? How might these challenges affect the conclusions that you might be able to draw from your observations and your ability to generalize from those conclusions?

Research Design Exercises

1. You are designing a study to evaluate the effectiveness of a new text-entry method for messaging on cell phones. Due to the popularity of messaging among college students, you decide that the undergraduate student body at your school would be an appropriate pool of potential participants. What would you want to know about the habits of these students regarding text messaging? You might be interested in comparing the performance of computer science students against students from other fields. Are there any other attributes of the students that might make for interesting comparisons? Given the male-female imbalance in computer science, what problems might this comparison involve?
2. Your research design for the study of text-entry on cell phones involves asking users to perform a set of tasks in a laboratory. As they will not be using their own phones, there is little, if any, privacy risk. What other risks might this study pose, and how would you inform users about them?
3. Find the website or other information about your institutional review board. Examine the policies and procedures specific to your institution, and write a draft informed consent form for the study described in Exercise 1.
4. Studies of how users respond to events that interrupt their work (Gluck, Bunt and McGrenere, 2007) present a challenge in design. If participants are told that the study is investigating reactions to

interruptions, they may be more sensitive to those events than they would otherwise be. A deceptive study, in which the subjects were provided with an alternative description of the goals of the study, might be one way to get around this problem. How might you describe a deceptive study for examining reactions to interruptions? How would you describe this study in an informed consent form? What would you discuss in the debriefing sessions?

References

- Azar, B. (2000) Online experiments: Ethically fair or foul? *Monitor on Psychology*, **31**(4).
- Bruckman, A.S. (2002) Ethical guidelines for research online. Retrieved July 2, 2007, from <http://www.cc.gatech.edu/~asb/ethics/>.
- Brush, A.J., Ames, M., and Davis, J. (2004) A comparison of synchronous remote and local usability studies for an expert interface. *Extended Abstracts of the ACM Conference on Human Factors in Computer Systems*, 1179–1182.
- Centers for Disease Control and Prevention (2007) *US Public Health Service Syphilis Study at Tuskegee*. Retrieved May 31, 2007, from <http://www.cdc.gov/nchstp/od/tuskegee/>.
- Cook, T.D. and Campbell, D.T. (1979) *Quasi-Experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Company.
- Davis, T.C., Holcombe, R.F., Berkel, H.J., et al. (1998) Informed consent for clinical trials: A comparative study of standard versus simplified forms. *Journal of the National Cancer Institute*, **90**(9):668–674.
- Finn, P. and Jakobsson, M. (2007) Designing ethical phishing experiments. *Technology and Society Magazine, IEEE*, **26**(1):46–58.
- Frankel, M.S. and Siang, S. (1999) Ethical and legal aspects of human subjects research on the Internet. American Association for the Advancement of Science. Retrieved July 2, 2007, from <http://www.aaas.org/spp/sfml/projects/intres/report.pdf>.
- Gluck, J., Bunt, A., and McGrenere, J. (2007) Matching attentional draw with utility in interruption. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 41–50.
- Haney, C., Banks, C., and Zimbardo, P. (1973) Interpersonal dynamics in a simulated prison. *International Journal of Criminology and Penology*, **1**: 69–97.
- Jones, J.H. (1993) *Bad Blood: The Tuskegee syphilis experiment*. New York: Free Press.
- Kraut, R., Olson, J., Banaji, M., et al. (2004) Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, **59**(2):105–117.

- Larson, K. and Czerwinski, M.** (1998) Web page design: Implications of memory, structure and scent for information retrieval. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 25–32.
- Lazar, J. and Preece, J.** (1999) Designing and implementing web-based surveys. *Journal of Computer Information Systems*, 39(4):63–67.
- Macefield, R.** (2007) Usability studies and the Hawthorne Effect. *Journal of Usability Studies*, 2(3):145–154.
- Maloney-Krichmar, D. and Preece, J.** (2005) A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer–Human Interaction*, 12(2):201–232.
- Meyer, J. and Bederson, B.** (1998) *Does a Sketchy Appearance Influence Drawing Behavior?* University of Maryland, Human–Computer Interaction Lab. Retrieved October 15, 2009 from <http://hcil.cs.umd.edu/trs/98-12/98-12.pdf>.
- Milgram, S.** (1963) Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67(4):371–378.
- National Cancer Institute** (2001) A guide to understanding informed consent. Retrieved June 4, 2007, from <http://www.cancer.gov/clinicaltrials/conducting/informed-consent-guide/>.
- National Commission** (1979) The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. Retrieved June 27, 2007, from <http://ohsr.od.nih.gov/guidelines/belmont.html>.
- Nielsen, J.** (1994) *Usability Engineering*. San Francisco: Morgan Kaufmann.
- Orne, M.T.** (1962) On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11):776–783.
- Patrick, A.** (2007a) Commentary on research on new security indicators. Retrieved June 22, 2007, from <http://www.andrewpatrick.ca/essays/commentary-on-research-on-new-security-indicators/>.
- Patrick, A.** (2007b) Privacy practices for HCI research. Retrieved June 4, 2007, from <http://www.andrewpatrick.ca/essays/privacy-practices-for-hci-research/>.
- Pellegrino, E.D.** (1997) The Nazi doctors and Nureinberg: Some moral lessons revisited. *Annals of Internal Medicine*, 127(4):307–308.
- Petrie, H., Hamilton, F., King, N., and Pavan, P.** (2006) Remote usability evaluations with disabled people. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1133–1141.
- Schechter, S., Dhamija, R., Ozment, A., and Fischer, A.** (2007) The emperor's new security indicators: An evaluation of website authentication and the effect of role playing on usability studies. *Proceedings of the IEEE Symposium on Security and Privacy*, 51–65.
- Sharma, K.** (2005) Can clinical trials ever be truly ethical? *The Hindu*.

Shavers, V.L., Lynch, C.F.L., and Burmeister, L.F. (2000) Knowledge of the Tuskegee study and its impact on the willingness to participate in medical research studies. *Journal of the National Medical Association*, **92**(12):563–572.

Shneiderman, B. and Plaisant, C. (2009) *Designing the User Interface*, 5th edition. Boston: Pearson Addison-Wesley.

Slater, M., Antley, A., Davison, A., et al. (2006) A virtual reprise of the Stanley Milgram obedience experiments. *PLoS ONE*, **1**(1).

Smith, S.S. and Richardson, D. (1983) Amelioration of deception and harm in psychological research: The important role of debriefing. *Journal of Personality and Social Psychology*, **44**(5):1075–1082.

Society for Research in Child Development (1991) Ethical standards for research with children. Retrieved June 22, 2007, from <http://www.srcd.org/ethicalstandards.html>.

Steriadis, C.E. and Constantinou, P. (2003) Designing human–computer interfaces for quadriplegic people. *ACM Transactions on Computer–Human Interaction*, **10**(2):87–118.

Varnhagen, C.K., Gushta, M., Daniels, J., et al. (2005) How informed is online informed consent? *Ethics & Behavior*, **15**(1):37–48.

Zimbardo, P. (2008) Stanford prison experiment. Retrieved June 2, 2008, from www.prisonexp.org.

15

Working with research participants with impairments

- 15.1 Introduction
- 15.2 How many participants?
- 15.3 Proxy users
- 15.4 Multi-Population Studies
- 15.5 Recruiting users through community partners
- 15.6 Pilot studies
- 15.7 Scheduling users with impairments
- 15.8 Documentation for users with impairments
- 15.9 Differing levels of ability
- 15.10 Bringing extra computer parts
- 15.11 Payment

15.1 Introduction

Chapter 14 talks about approaches for and issues that arise when working with human participants in research. As the number of research projects involving users with impairments grows, it is important also to examine the specific concepts, issues, and challenges of doing human-computer interaction (HCI) research with users with various impairments. Computer technology is now being used everywhere, by everyone, on a daily basis, for work, for pleasure, for communication, and for overall living. This includes users with perceptual impairments (e.g. hearing and visual), motor impairments (e.g. limited or no use of hands, arms, legs, or mouth) and cognitive impairments (whether lifelong impairments, such as Down Syndrome and autism, impairments that develop over time, such as dementia and Alzheimer's Disease, or event-based impairments, such as aphasia).

The grouping of "users with impairments" is itself somewhat artificial. It encompasses lots of different individuals with different impairments, abilities, and strengths; all they may have in common is that they have the label "impairment" or "disability" attached to them. For instance, individuals who are blind, and individuals who have Alzheimer's Disease may have practically nothing in common. And people that are often grouped together in research may be exact opposites. For instance, in evaluating technologies for people with cognitive impairment, some researchers have grouped together young adults with autism and Down Syndrome, when they are polar opposites in social skills, motor skills, and intellectual skills. This is important to remember: you can't just group together people with different impairments under that one large umbrella. While research on users with perceptual and motor impairments has existed since the 1970s, only recently have researchers tackled the challenges of designing computer interfaces for users with cognitive impairments (Lazar, 2007b) and only rarely have researchers worked with individuals with multiple impairments.

The goals of HCI research on users with impairments are the same as research with other users, to understand the phenomena surrounding computer interfaces and usage patterns. Because the users have a complex story, it is important to involve those individuals in HCI research, design, and evaluation. You can't just take guidelines from the research on interface design for people with impairments, and you can't just take proxy users that represent the users with impairments. You must work with users with impairments themselves. The overall research methods (experimental design, surveys, time diaries, case studies, etc.) are the same as for other users. However, the logistics of performing this type of research are what makes it different. There are differences in the number of participants, how you recruit participants, where you perform your research, how you get them to sign IRB forms, and how you pay users with impairments for their participation. These differences in logistics are covered in this chapter. Due to these complex logistics, it is realistic to say that it may take more time to do research involving participants with impairments. It is intensive, but you should do it anyway! In addition, some technologies that start out as assistive technology for a specific impairment population wind up later becoming popular among the general population. So, research that leads to improved interface and design experiences for people with impairments may eventually lead to interfaces that are better for the general population!

15.2 How many participants?

One of the greatest challenges of doing research with users with impairments is access to the participants themselves. Historically, many general research studies utilize computer users that are easy to gain access to. This includes students at universities, local business professionals, and children in schools (Lazar and Norcio, 2000). Finding appropriate users with the specific impairment that is the focus of the study can be a challenge. In doing research with the general population of users, it is often expected that a research study would have a minimum of 20–30 users, to be considered valid (see Chapters 2 and 3 for more information on sample sizes). These expectations may not be realistic for users with impairments, as it might be impossible to get access to so many users in one geographic area with a specific impairment.

The generally accepted approaches for dealing with the issue of access to appropriate participants for research focusing on users with impairments are small sample sizes, distributed research, and in-depth case studies. Choosing the most appropriate approach will depend on the nature of the research questions. For instance, controlled studies often use small sample sizes or in-depth case studies. Research of a more exploratory nature (with fewer controls) can use distributed research.

15.2.1 Small sample sizes

For research focusing on users with impairments, it is generally acceptable to have 5–10 users with a specific impairment take part in a study. This is due to a number of factors discussed later in the chapter. For example, in the recent proceedings of the ASSETS conference (well-accepted as a high-quality conference on this topic), most of the research studies in which blind users had to be physically present to take part in the research had 15 or fewer blind individuals taking part in the research. This means that if a classic experimental design is used, that there will often be no more than one control group and one treatment group, as the number of participants does not allow for empirical tests for multiple treatment groups (see Chapters 2 and 3 for more information on experimental design). However, this is fine, as research on users with impairments is often not of a traditional control-group-treatment-group nature; instead, it is often exploratory, a hybrid of quantitative and qualitative research, or primarily qualitative.

15.2.2 Distributed research

A different approach for users with impairments is to do distributed research, where the users do the research in their own home or office, without researchers present, and data is collected via time diaries, surveys, or keystroke logging. While this lowers the control that the researchers have over the study, it generally allows for higher numbers of users (100 users or more) to take part. In addition, a number of the challenges discussed later in the chapter (such as scheduling and transportation) may not be present for distributed research. To see an example of this, see the Time Diary to Study User Frustration sidebar in Section 6.1.

15.2.3 In-depth case studies

Yet another approach is to do in-depth case studies, in which fewer users (say, between three and 10) take part in a more intensive way. These studies might involve data collection over several days or users being trained, or longitudinal studies. This is most appropriate when data cannot be appropriately collected in a short amount of time (say, two to three hours). For instance, for many complex software applications or devices, users really do need a period of training, as well as time to familiarize themselves with the tool. So, a two-day period of research for each user can be seen as a minimum for a case study. Ideally, longitudinal studies would examine how users adapt to and utilize a new application over 3–6 months. For an example of a case study that included training, see the iSonic Evaluation Case Studies sidebar. Due to the complex nature of this software/hardware application, an in-depth case study was the most appropriate form of research.

Research
In
Practice

iSonic Evaluation Case Studies

A software tool called iSonic was developed to allow blind users to explore coordinated maps and tables, using sonification on the maps. The goal of this project was to create an accessible equivalent to information visualization for blind users, which would allow for coordinated data views using both tables and maps, along with the ability to filter and zoom in on items of interest.

In sonification, different non-textual tones represent different values. Specifically in iSonic, users could listen to a “map sweep”: the users would hear various tones to represent, for example, the population of various states in the US (or counties), starting from the northwest, crossing to the northeast, and then going from the southwest to the southeast. After the iSonic tool was developed, a series of case studies took place to evaluate the tool.

Seven blind users took part in the research study and three sets of data were used: one for training (data on the 50 states in the US), one for actual evaluation (data on the 24 counties of Maryland, where the evaluation study was taking place), and one for post-evaluation free exploration (data on the 44 counties of Idaho). For each user that took part, there were two separate sessions on two days. On the first day, the users interacted with a tutorial on iSonic and practiced using all of the features and sample tasks. On the second day, the user attempted a series of tasks, using both Excel and the iSonic tool, to compare the performance of those tools. For instance, these tasks included “Name the five counties [in Maryland] with the lowest housing unit value” and “What is the population of Dorchester County [Maryland]?” After the tasks were completed, there was a short period of interviewing users. Finally, the users were then encouraged to freely explore a new map (the map with data for Idaho). Between the seven users, a total of 42 hours of data was collected (Zhao *et al.*, 2008).

15.3 Proxy users

In the past, some researchers would use “proxy users”, where individuals without impairment would represent individuals with impairment during design or research. This could include people with no connection to the impairment and people with some knowledge of the impairment. Examples of people with no connection to the impairment include blindfolding people who can see or tying people’s hands behind their back to simulate users with motor impairments. These “simulations” are generally not encouraged for any type of research as, over time, users with perceptual or motor impairments learn to compensate by improving the use of their other senses or body parts. Someone who is blind has learned to rely more on their hearing than someone who can see. Even if the users of interest and users without any impairments are considered to have equal skill in some area (for instance, good quality speech), the impairment makes users perceive the technology differently. So, it is inappropriate to test speech-recognition solutions for users with spinal cord injuries, by using users without any impairment, based on the claim that they have similar quality speech (Feng, Sears and Law, 2005). Since users often compare a new technology to a previously used technology or option, the comparisons are very different.

There are some situations where it is appropriate to use people who are familiar with the users and impairments to represent the users themselves. These are generally situations where users are unable to communicate, or are unable to process information due to their impairment. For instance, one study used speech–language pathologists who work closely with individuals with aphasia, instead of the actual users themselves, to get an understanding of user needs (Boyd-Graber *et al.*, 2006). In another study, caregivers and family members were used as the primary information sources for designing technology for individuals with Alzheimer’s Disease (Cohene, Baecker and Marziali, 2005). In both of these cases, the users of interest were themselves unable to communicate. In another study, parents answered questions about the computer usage of their children with Down Syndrome (see sidebar).



Children with Down Syndrome

There has been almost no research into the computer usage of children and young adults with Down Syndrome. Two of the co-authors of this book created a survey study, to learn more about the computer usage patterns of children and young adults with Down Syndrome.

The goal was to establish some baseline data about how children with Down Syndrome used computers, and what challenges they faced. Since the project was geared towards individuals with Down Syndrome between the ages of five and 21, asking the children and young adults themselves to fill out the survey would not have

been feasible. While the older individuals (teenagers and up) might have been capable of filling out the survey, they might not have had the level of reflection and language required to understand and explain exactly what they do on the computer. Certainly, the younger children would not have been able to respond to the survey.

Furthermore, since most participants were under 18 years old, their parents would have been required to provide the informed consent to participate. Therefore, parents, as individuals who could give consent for participation and were most familiar with the computer habits and skills of their children, were considered appropriate proxies for their children with Down Syndrome (Feng *et al.*, 2008).

Note that, even with cognitive or motor impairment, many users can communicate by using some form of assistive and augmentative communication (AAC) device. You should never use proxy users when users can communicate but the researchers don't speak their language (such as people who are deaf and use sign language or blind-deaf users who use Braille or finger-spelling). In those cases, you need to access individuals who can communicate and translate with the users in their own language.

Another situation where proxy users might be appropriate is when a specific application or tool is being developed and it is undergoing multiple iterations before a proof-of-concept is complete. If users with the specific impairment would not be available to take part in all stages and all iterations of design, then proxy users might be suitable in limited stages and limited circumstances, for testing purposes. However, they should closely be followed up by evaluations with users who actually do have the impairment.

15.4 Multi-Population Studies

Given that users with impairments are really a mosaic of different communities with different needs, it is sometimes important to test an interface with either multiple impairment groups, or a combination of impairment groups and users without impairment. There are generally two approaches for developing interfaces for users with impairments (Lazar, 2007a):

- Try to make an interface (for a website, digital library, or operating system) that works well for a majority of users with impairments, especially perceptual and motor impairments. Usually, this is the scenario where the users have the same end task goal as users without impairments (such as accessing an article or purchasing a song online), and they are simply utilizing alternative input or output devices (Slatin and Rush, 2003).
- Design an interface that is optimized for a specific user group. This is the approach that tends to be used for people with severe cognitive impairment, including children with autism and adults with Alzheimer's Disease or aphasia (Cohene, Baecker and Marziali, 2005; Moffatt *et al.*, 2004; Tartaro, 2007). The needs of the population are so specific,

that the interface, and the corresponding task scenarios and applications, are so focused on the specific needs of the user population that they are unlikely to meet the need of other populations.

For the first approach, interfaces are generally designed for a combination of the general user population without impairments and a few targeted user groups (such as users with hearing impairment, visual impairment, and spinal cord injuries). In these cases, it is generally important to make sure that an interface that is easy to use for users without impairments is also easy to use for users with certain impairments. Therefore, it is necessary to test the interfaces with two or three different user groups (users without impairments and the targeted two user groups). While we often talk about the goal of universal usability, the reality is that you can never test an application or interface with every possible existing user population. Often, an application is labeled “universally usable” when it is evaluated with three or four user populations.

It might make sense, for instance, to test a new form of CAPTCHA (a web-based security tool to differentiate between a software program and a human being) using both blind users and users without any impairments. While a CAPTCHA that works for blind users is nice, it will only be used in practice if it is also easy to use for the typical user without impairments (Holman *et al.*, 2007). In reality, if an interface works well for users with perceptual or motor impairments, that’s wonderful, but companies and organizations will not implement those interfaces if they in any way degrade the user experience for users without impairments. In these situations, where an interface must be easy to use for users with and without impairments, multi-population studies are needed to involve both the general user population and a few selected impairment populations. In those cases, the general user population is NOT serving as a proxy user, but rather, is part of the targeted user population for an interface.

15.5 Recruiting users through community partners

At this point, it should be clear that recruiting actual users with impairments is necessary for all forms of research, including usability testing. The next question is, how is this done? You can’t just place signs in the computer department or on campus saying, “we want users with spinal cord injuries to take part in our research study,” as there are often not a sufficient number of individuals with impairments on university campuses. The target population may not see the signs (if they have a visual impairment) or have access to the spaces where the signs are posted. The best way to recruit users is usually to partner with a community-based group that focuses on the impairment of interest to the research. Most people with impairments have some sort of organization, support group, or coordination point. For instance, there are organizations for people with visual and hearing impairment, organizations for people with spinal cord injuries, and organizations for people with Alzheimer’s Disease. In cases where the impairment impacts on the ability to live an independent life, these organizations

often include caregivers and family members. Many university campuses have an office that provides support to people with impairments.

It is usually good to approach these organizations for help in recruiting users. However, simply saying, "we want to do some research, and we need your help in recruiting users" is not sufficient, and it is hard to establish immediate trust (Feng, Sears and Law, 2005). If you really care about these user populations, then you need to become involved with the community-based group for the long term. Most of these organizations get multiple requests for help, and they may be leery of "drive-by research," where you ask for their help, do the research, and then never show up or contact them again.

Some organizations are geographically based and you may want to contact their national offices. For instance, the Royal National Institute of Blind People in the United Kingdom and the National Federation of the Blind in the United States are leading organizations for blind individuals. Similarly, there are many other groups, such as the National Down Syndrome Congress in the United States and the Down Syndrome Association in the United Kingdom. While national organizations are common, other organizations may work at the grassroots, with local city-based groups that do not coordinate with each other. If possible, you should become a part of these organizations: go to their meetings, meet people, get involved in their community, and take part in fundraisers. If there is a regional or national convention, it is important to attend that gathering. At these gatherings, it is possible to better understand the logistics and challenges involved for that population, which can help with the planned research in the future. But it isn't sufficient to go to the meetings just to learn about issues such as Braille handouts or physical room limitations for individuals in wheelchairs. The end goal should not simply be to further your research, but to further the cause of these individuals and their quality of life. Your research is simply a piece of that long-term goal. As such, your partnership needs to be a two-way street. If you are asking for their help, then they should be able to expect your help. You should find a way to compensate the organization for their assistance to you. When your research is complete, you should make sure that the organization receives copies of any final reports.

Rehabilitation centers that are often sponsored by local governments or industry provide training and modifications to help adults with certain impairments move into the workforce. These organizations can often be sources of participants for research.

If you are working with a community-based organization that specializes in a certain impairment, the goal of your research is to further their cause and improve the quality of life for individuals with the specific impairment by improving understanding of HCI issues for the user population. If the only goal you have is to further your own professional career, with little concern for the needs of the population, look elsewhere. Working with users with impairments is a long-term, emotional, involved process, with great societal benefit and long-term payoffs in the quality of life for individuals. Expect that the organizations involved will come to count on you and consider you a part of their cause. Invest in the long term or get out of the game. End of sermon.

When recruiting users, it is important to understand their preferred method of communication and any related challenges. For instance, e-mail may not be the preferred option for users with spinal cord injuries, as it may be harder for users with SCI to generate text. Instead, phone calls might be the preferred option (Feng, Sears and Law, 2005). Obviously, phone calls might not work well for deaf people, who may prefer e-mail or text messaging. Other user populations may have different challenges in communication. For instance, e-mail is often a preferred method of communication for blind users. However, due to the large amount of time required for them to process spam e-mail, blind users tend to have very strong filtering on their e-mail. E-mails sent to multiple blind users using the BCC option will not make it through the spam filter to most users (Lazar *et al.*, 2005). So for blind users, it is important to place the recipients' e-mail address in the To line, not in the CC line or BCC lines. Another approach might be to use the phone, but it is important to ask permission in advance before doing so. For users with some types of cognitive impairment, it may be necessary to contact caregivers.

15.6 Pilot studies

Due to the logistics involved, it is often very necessary to do pilot studies before beginning any real data collection. Your simulation in the lab, or your expectations of how a user will interact, are likely to be very different from the reality. While this is true in any type of HCI research, it is especially true in working with users with impairments. Since you may have access to a limited number of users and you won't have any opportunity to do the data collection a second time, you need to confirm or address your perceptions early on in the process by doing a pilot study with one or two users.

Pilot studies can uncover a number of problems. For instance, is the documentation accessible for the specific user population? Users with spinal cord injuries can't physically handle documentation, and blind users may not be able to use printed materials or even Braille materials (approximately 10–20% of blind individuals are fluent in Braille). Users in wheelchairs will need physical settings, including computer desks, that can accommodate their wheelchairs. Other technical problems may arise. For instance, any text documents for blind users must work under multiple screen readers (Window-Eyes and JAWS), multiple operating systems (OS X, Win XP), and multiple text editors (MS-Word, Word Perfect, and Notepad), as well as various combinations of screen reader, operating system and text editor (Lazar *et al.*, 2005). Sometimes the file format that works best is Rich Text Format, which tends to work with most text editors. In doing a pilot study, you may find out that the participants expect to use aids (such as a portable notetaker, voice recorder, or electronic device) or expect you to have aids available to them (Sauer *et al.*, 2009). Generally, you need to be aware if all participants are using certain aids; if only some of them do, you need to find a way to compensate for that in your data collection.

One or two users in the pilot study are generally enough, just to confirm that you are on the right track and that there are no major problems with logistics. If you have worked

with a specific user population for a long time, you may have a few users that you collaborate with regularly, who are comfortable with you, and are willing to help you test out materials and serve as your “reality check.” Whatever flaws or problems are discovered during the pilot study should be modified and accounted for before the main study begins.

15.7 Scheduling users with impairments

It is important to remember that users with perceptual, cognitive, or motor impairments frequently do not drive a car. They may rely on rides from others, public transportation, taxis, and scheduled services to get from point A to point B. Therefore, these users must typically be scheduled in advance. It is often not possible for these participants to make transportation plans, or change them, at the last minute.

Rather than asking participants to come to a university or remote location, it is far better for researchers to offer to go to a home or workplace location. To help ensure the safety and security of researchers entering participant homes, it is preferable to go in teams of at least two researchers. By visiting users in their home or workplace, it alleviates the need for the user with an impairment to schedule transportation to a new location. In addition, getting a glimpse of the user in their own environment, using their own technical setup, is likely to lead to a more ecologically valid data collection effort. Many blind users have a screen reader (such as JAWS or Window-Eyes) customized to their specific needs. The speed of speech, how links are read, and even the type of voice (e.g. American English vs. British English) are personal preferences that may not be obvious to the researcher, but can be very important to the user. Visiting the user in their natural environment allows the user to be most relaxed and productive and yields the most ecologically valid data.

The major drawback of visiting users in their work or home environment is that you tend to have less control over the environment (Feng, Sears and Law, 2005). If users are able to come to a research lab, this offers the researchers more control over the layout and noise in the environment. However, aside from the transportation challenge, there is another major challenge: the accessibility of the building. Researchers must be completely certain that the building that they expect users to come to is accessible. This means that the doors must be wide enough, restrooms must have accessible stalls, elevators must be present, and Braille must be available on all signs. In addition, some users may have service animals working with them (Feng, Sears and Law, 2005).

It is also important to note that a large number of users with an impairment do not have a job; those who are employed may be very sensitive about missing work for an outside research project. They are unlikely to let a research study interfere with their job performance (Lazar, Feng and Allen, 2006). So when possible, visit them on-site, either at their workplace or their home (which is sometimes preferable because it won't interfere with work). Also, note that it may be necessary to schedule research sessions during evenings or weekends.

It is important for researchers to understand that the variety of users and the various levels of severity of the impairment (see Section 15.9) mean that the time involved for a user

to take part in a research study might be relatively unpredictable. The researcher's schedule should be left flexible enough that it is not a problem if a user takes much longer for data collection than is expected. In addition, many users with impairments are determined to prove that they can accomplish tasks. This means that if the time period is limited for the specific user's data collection, they may still want to continue and feel the need to complete the task. For a researcher to tell the user that "time is up" may be met with resistance. This is not generally a problem, except that it needs to be accounted for in the scheduling of users.

15.8 Documentation for users with impairments

Often, there are a number of documents that are required for participation in a research study. These include human subjects forms (also known as institutional review board (IRB) forms – see Chapter 14 for more information), instructions, task lists, and questionnaires. In traditional paper format, these forms may pose a problem for users that are print-disabled (blind or with low vision or dyslexia) or that can read but may have problems handling forms (such as users with spinal cord injuries). It's also important to note that in some cases, if children with impairments are involved in the research, then the researchers themselves may be required to submit their own approval paperwork related to criminal record background checks.

15.8.1 Human subjects forms

To start with, nearly all research projects involving humans require that participants be informed of their rights, and this usually takes the form of a human subjects or IRB form (see Chapter 14 for more information on what rights human participants have). Most human subjects forms require handwritten signatures, as per university or institutional requirement. This may be a problem for a number of user populations.

Users with motor impairments, especially those that are unable to use their arms, may not be able to use a pencil to sign a form or handle a form. An audio recording, or a video of the user, agreeing to take part in the study, hopefully will be acceptable to the institutional review board. For users with certain types of cognitive impairment, it's questionable whether they would be able to sign a legal document. A caregiver, who has legal standing, might need to provide the signature. For children with an impairment, often the parents need to give their approval for participation in the research project. Blind users may be able to sign the form, however, they will need guidance on where to sign the form (in addition, it's questionable whether we should ask participants to sign a form that they cannot read first). For users that either cannot read or handle the form, it is good practice to send an electronic version of the form beforehand, so that the user can read and be comfortable with it. Be sure to understand the specific policies relating to IRB forms from the organization that approved the research study (usually a university). For instance, many universities accept nothing but a signed, paper-based form. Some universities are beginning to accept electronic versions of informed consent (see Chapter 14). It is helpful to check if your institutional review board

can accept some modified form of informed consent, which is more appropriate to the user population. If the institutional review board will not accept audio or video recording of a user giving consent, there are work-arounds that can be utilized.

If a sponsoring organization requires signed forms from blind users, there are two popular ways of guiding blind users to the appropriate place to sign on the form. One method is to provide a signature guide (a small piece of plastic with a window in the middle, to indicate where the signature should be – see Figure 15.1). The other method is to attach a Braille label right below the signature line. The Braille label could say something along the lines of “sign above” (Lazar *et al.*, 2005). While this might not be meaningful for the majority of blind individuals who are not able to read Braille, the tactile information provided by the top line of the label can provide useful information on where the signature should be placed. Careful attention to details such as these can help build trust and confidence with participants, as they may appreciate that you’ve made the effort to make things work for them.

The discussion of blind users and Braille brings up another important issue. You must be aware of the diversity within user populations. For instance, print forms are relatively useless for blind users, unless they have a scanner available. So it might seem that forms in Braille would be the appropriate alternative. You first need to check that all of your user population can read Braille and that your university or other sponsoring organization will accept forms written in Braille. Most individuals who are blind cannot read Braille; if you want to test screen reader usage, by having forms printed in Braille, you would limit yourself to the estimated 10–20% of blind individuals that are fluent in Braille (National Federation of the Blind, 2006). And, of course, you would also need access to some form of Braille printer to print good-quality Braille. If you are not very familiar with the characteristics of a certain population and are working with them for the first time, you really should get advice on all of your research plans from someone who has years of experience working with that user population.

15.8.2 Research documentation

Once the issue of human subjects forms has been addressed, there are issues surrounding the other documentation in the research study. For instance, participants in research studies must often either read material, or record their responses, on paper. If users are unable to read printed documents or have trouble handling physical documents, then there are other options. One option is to provide all of the materials in electronic format, which can be used both for reading and for recording responses. Plain text versions of all documentation can be made available to the users, at the time of the research study. Only the IRB form should be made available beforehand, as providing actual study documents could lead to learning effects. Electronic forms introduce another complicating factor into the research study. For instance, what happens if some users are more experienced with text readers or word processors than other users? Will that difference, even though it is not being measured or controlled for, make a difference in the outcome of the research?

INFORMED CONSENT FORM FOR THE RESEARCH EXPERIMENT

Purpose of the Project:

Dr. Jonathan Lazar and his students are creating a research study to learn more about how blind users using screen readers become frustrated while surfing the web. With a better understanding of what frustrates users, we can come up with ways to improve the user experience. We hope that the results of this study will have beneficial effects to make computers less frustrating.

Procedures for Participants:

You will be asked to fill out a pre-session survey. After filling out the survey, you will be asked to perform your normal computer tasks for a minimum of two hours. Whenever you feel frustrated, you are asked to fill out a form documenting your frustrating experience. After performing your normal tasks for a minimum of two hours, you are asked to fill out a post-session survey. You should then mail all documents back to Dr. Lazar at Towson University.

Confidentiality:

Participation in this study is voluntary. All information will remain strictly confidential. Although the descriptions and findings may be published, at no time will your name or any other identification be used. You are at liberty to withdraw your consent to the experiment and discontinue participation at any time without prejudice. If you have any questions after today, please contact Dr. Jonathan Lazar at 410-704-2255 or contact Dr. Patricia Alt, Chairperson of the Institutional Review Board for the Protection of Human Participants at Towson University at (410) 704-2236.

I have read this form and had all of my

Subject's Signature _____

Date _____

Figure 15.1 An IRB form with a signature guide for blind users.

The other option is to verbally instruct the user on what to do and ask them to respond verbally. While this is very appropriate, the major caveat here is to make sure that rules are created to guide the researchers on what they do and do not say. For instance, is there a limit on the number of times that the researcher can repeat instructions? Do the researchers refuse to answer questions outside the scope of the instructions? Can they spell out words? For instance, if the research study was investigating web searching habits, it would not be appropriate for the researchers to give hints or provide guidance to the users. Therefore, there should be clear rules for the researchers on what they can and cannot say, so that there is consistency across all users taking part in the research study. Obviously, you must tailor the documentation to the needs of the participants. For instance, verbal instructions would not work well for users with hearing impairment or who are deaf. If those users have vision, then paper documents may be preferred. But if users have a motor impairment, such as a spinal cord injury, in which case handling documents and recording responses on paper might be problematic, then audio recording might be a good option. If users are deaf-blind, Braille may be the preferred option. As always, you must know your participant population very well.

15.9 Differing levels of ability

Ability levels may vary widely among users with a specific impairment (Jaeger, 2009). Assumptions should never be made, for instance, about “what users with aphasia are capable of.” Since many impairments are due to underlying medical or health causes, the severity of the impairment will vary among different users. Most impairments are not binary, that you either have them or don’t. People can have partial impairments (such as partial hearing or visual impairment). People can have varying severity of impact (for instance, mild, moderate, or severe aphasia, Alzheimer’s Disease, or dementia). Even impairments that at first seem to be very clear and binary are not. For instance, there are different types of amnesia, based on what type of memory capability has been lost. While trisomy Down Syndrome is the most common form, there is another type of Down Syndrome, called “mosaic Down Syndrome”, that is much rarer, but generally has a lower impact on cognitive performance. In all of these situations it is important to fully understand the nature of the population, by consulting with experts in that specific impairment. In addition, standardized tests that measure the severity of the impairment can be very useful, as long as they are properly conducted and interpreted (Moffatt *et al.*, 2004).

Not only does the severity of the impairment influence interface design, but even for people at the same level of impairment, there are a number of other factors that influence performance on interface-related tasks, including: confidence, self-efficacy, and previous experience with using computers. The results are not always what they seem and it takes a lot of experience with a specific user population to understand this.

For instance, research tasks that might take user A only one hour might take user B 3.5 hours. In a typical population without impairments, this would lead the researcher to believe

that either users B's performance is lower, or maybe there is a problem with the equipment that user B is utilizing (e.g. it is older equipment or network connections). However, this would not necessarily hold true for populations with disabilities. For instance, newer users of a certain application or tool (such as head tracking) might be satisfied with completing a series of tasks in 3.5 hours. This same amount of time might be frustrating to someone who has utilized the equipment for years. Each user with an impairment (or a combination of impairments) is a unique individual, with a unique performance speed that they alone consider to be their average "default speed". The "default speed" should be taken into consideration to determine individual usability. However, the "default speed" can also be a complication when trying to compare the performance of a group of users with a specific impairment. For instance, typical data input and output speeds vary more greatly for users with impairments than for the general user population. As an example, blind users listen to their screen readers at varying rates, and tend to think that any speed that is not their pre-set speed is either too fast or too slow. Often in studies with blind users, you want to remove the potential confounding factor of having various screen reader speeds in the mix by using one screen reader speed for every participant. This is a very good thing for your research design, but may frustrate the individuals who participate.

In another example of the complexity of user differences within a specific impairment population, for a screen reader user who listens to JAWS at a very rapid rate, they may be frustrated if a task takes more than five minutes to complete. Another user, who listens to the screen reader at a much slower speed, may be very satisfied if the same task takes 20 minutes to complete. Their personal expectations of performance may not always be obvious to the researcher and this may be hard to measure. Experience with the computer and confidence may also play a role. For instance, imagine three blind users, all of whom are attempting the same task. User A may give up after two minutes of attempting the task, because they know that they typically can only find information using four different navigation methods, and once they have attempted all four navigation methods, it is pointless to continue, as they are confident that they would not be able to use any other method and succeed. User B may also give up after two minutes, but because they have low confidence. They are not confident in their abilities and think it is unlikely that they will be able to complete a task. User C does not give up, even after 45 minutes of attempting a task. While the computing skill set of user C might be high or low, they are confident in their abilities, and they repeatedly say, "I am not a quitter. I will keep going until I am able to complete the task." In this example, time is not directly correlated to experience or confidence, but rather, is influenced by both. The authors of this book have personally witnessed all three behaviors.

15.10 Bringing extra computer parts

When visiting users with impairments in their home or workplace, it's important to understand that their setup may not be what most researchers are used to, and that technical setup will be out of the researcher's control. For instance, blind users may not have a working

monitor, deaf users may not have working speakers, and users with motor impairments may not have a working mouse. Since many of these users have purchased a “standard package” of CPU, monitor, and peripherals from a computer company, if pieces of hardware that are useless to them break, there is no real incentive for the users to replace them. However, researchers often rely on these tools to understand the user interaction. For instance, often researchers who are visual will need to see the screen to understand what the screen reader is reading. If this is the case, you need to carry extra computer parts in your car when you visit the users. For instance, bring a monitor with you if you are visiting blind users in their workplace or home. Also bring standard cables (such as video and USB cables). If doing multiple on-site visits, it is good practice to take extra parts (monitors, cables, speakers, mice, external keyboards) with you at all times, and simply leave them in the car, as you never know when you may need them.

A circular icon with the text "Research in Practice" inside, connected to the main text by a black line.

Participants Getting Frustrated

What happens if a user with an impairment is taking part in a study, is not successful at completing any of the tasks, and is getting frustrated? This person is getting agitated, is still trying to complete the tasks, but clearly is not making any progress. What happens next? This is a realistic question.

For the researcher who is monitoring this user, it is an upsetting time. Although our research studies in HCI typically do not endanger health or leave lasting emotional effects, it is certainly possible that a situation of this nature could occur which could leave the user angry and upset. Aside from a few rare studies designed to frustrate people on purpose, such as (Riseberg *et al.*, 1998), HCI research is generally not designed to aggravate the user.

There are a few options. The researcher can remind the user that they have the right to end their participation in the experiment, at any time, with no adverse consequences (which is typically a standard requirement in IRB forms). As part of this reminder, the researcher should note that whatever payment is due to the user for participation will be given to the user, regardless of when they end their participation. But if the user does not want to end the session, what happens next? Perhaps the user can be offered a short break or a period of rest, which would allow them a few minutes to calm down. The researcher technically has the right to end the experiment if they feel that someone is starting to be harmed. However, for the researcher to unilaterally end the participation of the user also sets some bad precedents. If researchers frequently end user participation, there could be some bias injected into the research study. This is a tricky situation. Especially when working with users with impairments, who are often hard to recruit and replace.

15.11 Payment

When paying users for taking part in research, it is important to make sure that the form of payment will be useful to the users. For instance, gift cards for a specific store (such as a local bookstore) may not be useful for some people if they cannot use standard print materials. Also, gift cards that only work at a certain store may not be useful, if transportation is required to visit the store and use the gift cards. Gifts that are typically used to recruit university students for research, such as iPods, may also not be appropriate, as many users with impairments have very specific technical needs and may not want to use new devices. The best forms of payment are either cash or cash equivalents, such as cash cards. If those are not viable options, then at least a gift card should be given at a store that has online ordering options and an accessible website (such as Amazon) or that has many local branches and many types of merchandise. It is also important to note that users with impairments are typically paid more than users without impairments for their participation in HCI research.

Summary

Research involving participants with impairments can be challenging but it offers many rewards. The computer usage of many of these users has not been explored in as much depth as with the general population of users, so there are many great research questions that remain unexamined. And these topics need attention! With appropriate planning and attention to logistics, HCI research involving users with impairments can be very successful.

Review Questions

1. What are the three generally accepted approaches for dealing with the challenge of access to participants with a certain impairment?
2. What is an advantage of doing distributed research with participants with impairments and what is a disadvantage?
3. What is a proxy user? Why is the use of proxy users in research discouraged? What are the rare circumstances in which proxy users would be acceptable?
4. What are the two general approaches for developing computer interfaces for users with impairments?
5. Why is it important to make sure that an interface designed for a user with a perceptual or motor impairment is also maximized for the general user population?
6. What are some good places to look for potential participants with impairments?
7. Is e-mail always the best way to contact potential participants with impairments?
8. Pilot studies can be helpful in identifying potential challenges in logistics. Name at least three logistical challenges that a pilot study can uncover.

9. Why is transportation a challenge to participation for users with impairments and how can researchers address that?
10. What is a major challenge in using IRB forms for blind users? What about for users that are paralyzed?
11. Why are standardized tests of the severity of impairment useful in research studies?

Research Design Exercise

Imagine a research study that involves users who have both slurred speech and severe arthritis. These users do not have any cognitive impairment and their vision is average. The goal of this research study is to examine various input devices and determine which one is most effective for this user population. As hearing and vision is intact for most of these users, output is not a problem, only input is a problem. What might the transportation issues be for this population? What might the scheduling issues be? What might be the best way to communicate with these users? Would it be better to go out to their homes or have them come to the research lab at the university? How would you handle IRB forms? How would you give them the documentation on the tasks to be performed? How would you have them record responses (since, due to the arthritis, they may have trouble with writing)? How might these users like to be paid for their participation?

References

- Boyd-Graber, J.L., Nikolova, S.S., Moffatt, K.A., et al.** (2006) Participatory design with proxies: Developing a desktop-PDA system to support people with aphasia. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 151–160.
- Cohene, T., Baecker, R., and Marziali, E.** (2005) Designing interactive life story multimedia for a family affected by Alzheimer's Disease: A case study. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1300–1303.
- Feng, J., Lazar, J., Kumin, L., and Ozok, A.** (2008) Computer usage and computer-related behavior of young individuals with Down Syndrome. *Proceedings of the ACM Conference on Assistive Technology (ASSETS)*, 35–42.
- Feng, J., Sears, A., and Law, C.** (2005) Conducting empirical experiments involving participants with spinal cord injuries. *Proceedings of the Universal Access in Human-Computer Interaction Conference*.
- Holman, J., Lazar, J., Feng, J., and D'Arcy, J.** (2007) Developing usable CAPTCHAs for blind users. *Proceedings of the ACM Conference on Assistive Technologies (ASSETS)*, 245–246.

- Jaeger, P.** (2009) Persons with disabilities and intergenerational universal usability. *interactions*, 16(3):66–67.
- Lazar, J.** (2007a) Introduction to universal usability. In J. Lazar (ed.), *Universal Usability: Designing computer interfaces for diverse user populations*, 1–12. Chichester, UK: John Wiley & Sons
- Lazar, J.** (ed.) (2007b) *Universal Usability: Designing Computer Interfaces for Diverse User Populations*. Chichester, UK: John Wiley & Sons.
- Lazar, J. and Norcio, A.** (2000) Service research: Community partnerships for research and training. *Journal of Informatics Education and Research*, 2(3):21–25.
- Lazar, J., Allen, A., Kleinman, J., and Lawrence, J.** (2005) Methodological issues in using time diaries to collect frustration data from blind computer users. *Proceedings of the 11th International Conference on Human-Computer Interaction (HCI)* (on CD-ROM).
- Lazar, J., Feng, J., and Allen, A.** (2006) Determining the impact of computer frustration on the mood of blind users browsing the web. *Proceedings of the ACM Conference on Assistive Technology (ASSETS)*, 149–156.
- Moffatt, K., McGrenere, J., Purves, B., and Klawe, M.** (2004) The participatory design of a sound and image enhanced daily planner for people with aphasia. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 407–414.
- National Federation of the Blind** (2006) *Braille Usage: Perspectives of legally blind adults and policy implications for school administrators*. Baltimore, MD: National Federation of the Blind.
- Riseberg, J., Klein, J., Fernandez, R., and Picard, R.** (1998) Frustrating the user on purpose: Using biosignals in a pilot study to detect the user's emotional state. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 227–228.
- Sauer, G., Holman, J., Lazar, J., Hochheiser, H., and Feng, J.** (2009) Accessible privacy and security: A universally usable human-interaction proof. *Universal Access in the Information Society*. In press.
- Slatin, J. and Rush, S.** (2003) *Maximum Accessibility*. New York: Addison-Wesley.
- Tartaro, A.** (2007) Authorable virtual peers for children with autism. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1677–1680.
- Zhao, H., Plaisant, C., Shneiderman, B., and Lazar, J.** (2008) Data sonification for users with visual impairment: A case study with geo-referenced data. *ACM Transactions on Human-Computer Interaction*, 15(1):1–28.