16

Informed Web Surfing

The Social Context of User Sophistication

Eszter Hargittai
Northwestern University

There are many studies that look at how people use the internet (for a review, see DiMaggio, Hargittai, Neuman, & Robinson, 2001) and, in particular, what types of content they view online (e.g., Howard, Rainie, & Jones, 2001). There is a separate body of literature that looks at how people use information retrieval systems and, in particular, how people search for information on the Web (for a review of this literature, see Jansen & Pooch, 2001). However, these two areas of inquiry rarely intersect, leaving the discussion of what people do online in isolation from studies of what people are *able* to do online.

In this chapter, I report findings from a project that explores people's skills in locating content on the Web. Documenting differences in people's Web use

AUTHOR'S NOTE: I thank Paul DiMaggio for his insightful comments throughout this project, Stan Katz for his ongoing support, and Erica Field and Hank Farber for helpful discussions. I am also grateful to Inna Barmash, James Chu, Jeremy Davis-Turak, Edward Freeland, Anne Healy, Veronica McRipley, Carolyn Mordas, and Katy Niner for their assistance with various components of the project. I also express my gratitude to the many people who took time from their busy schedules to participate in the study. Generous support from the Markle Foundation is kindly acknowledged. The project has also been supported in part by National Science Foundation Grants SES9819907 and ITR0086143, a grant from the Russell Sage Foundation, and a grant from the Pew Charitable Trusts to the Center for Arts and Cultural Policy Studies at Princeton University. I am also grateful to the Fellowship of Woodrow Wilson Scholars at Princeton.

skills allows us to distinguish between how various kinds of people take advantage of the medium in different ways resulting not merely from their particular interests but also from their abilities in using the medium in ways that they prefer. This chapter focuses on internet users' skills and how people's online skills are embedded in a web of social relations. Skill, in this context, is defined as *the ability to find information on the Web effectively and efficiently.* Social relations refers both to people's immediate living partners (e.g., family and roommates) and to their wider networks of family, friends, and colleagues both physically close and distant.

First, I review literature on various people's use of the Web and the relationship between technology uses and people's social ties. Next, I outline the methods used in the project for studying how people locate content online, the sampling methodology, and how the data were analyzed. Then, I describe the findings from 66 interviews and in-person observations conducted during the summer and fall of 2001 in a New Jersey county. I present evidence to show that there is considerable variation in people's ability to find information online, and I explore what explains these differences. I pay particular attention to the relationship between people's online ability and their social relationships because these seem to play an important role in how people navigate the contents of the Web.

Web Use in Social Context

Many social factors may influence one's level of Web use sophistication. Although there is little literature on the social predictors of Web use skills, I draw on analogous work from the literature about how people use the medium to hypothesize about the role of social factors in people's online skills. Age, gender, education, and income all are factors that have been associated with the extent of people's connectivity and Web use patterns during the period of mass internet diffusion in the United States (National Telecommunications and Information Administration [NTIA], 1995, 1998, 1999, 2000, 2002) and are likely to be associated with people's online skills as well.[1] Younger generations have grown up with computers—if not in their homes, then often in their schools—leading to wider exposure to the technology among the young. Loges and Jung (2001) also found that older users engage in fewer activities online than do their younger counterparts.

Although the most recent figures show that men and women are network connected in about equal numbers (NTIA, 2002), Howard and colleagues (2001) did find differences in the types of activities in which female and male users engage while online. Men spend more time browsing for fun, looking up financial and product information, and consuming news online, whereas women spend more time looking up health information. Gender differences in online skills may be due to women having less leisure time at home (Kelsey, 2002; Lally, 2002). Insofar as home use is the most autonomous use in that it allows for the most freedom in exploring sites of interest to users, having less free time in the home to browse the Web may have a negative influence on women's Web use skills.

People's levels of education and income have been consistently related to their level of internet access and use and are likely to affect their level of Web use skills as well. Universities were the first to embrace the technology; thus, those who attended college during the past decade would have had exposure to the medium during their schooling. Moreover, people with higher levels of education are likely to have had more exposure to computer technology in general, and familiarity with the technology is an important first step in gaining access to the internet. Similarly, more affluent households are likely to have better resources, allowing them to purchase more up-to-date computers and high-speed internet access. With the spread of high-volume graphics and video technology, the need for various plug-ins and additional software to view many sites and documents (e.g., portable file documents [pdf]), high-capacity machines and high-speed connections are increasingly necessary to benefit from all the Web has to offer.

Of particular interest here is how people's social relationships also may be associated with users' online abilities. The literature on the diffusion of innovations emphasizes the importance of social support networks in the spread of new technologies (Rogers, 1995). Those with exposure to innovations in their surroundings are more likely to adopt new technologies such as personal computers (Dutton, Rogers, & Jun, 1987; Rogers, 1985). The availability of friends and family who are also internet users provides support for problems encountered while using the medium and is also a source of new knowledge via advice and recommendations.

For online skills in particular, this implies that people who are able to draw on their networks for information on how to use the medium will learn more quickly and will be exposed to a broader repertoire of online services than will those who have few people to whom they can turn for advice with their Web use. A study of home computer diffusion found that people were more likely to give up using the technology when they had no neighbors or friends to call on for support (Murdock, Hartmann, & Gray, 1992). By contrast, people whose social circles

include users knowledgeable about the Web can draw on their networks for site recommendations and suggestions when they run into problems.

The presence of children in the household may have particular effects on people's Web use skills. Controlling for time spent online (a factor likely influenced by the time spent on caring for children), presence of offspring may improve one's level of Web use sophistication due to knowledge passed on to parents by their children. A study of requests for help with internet use in 93 families found that children were significantly more likely to help adults in their home internet use than vice versa (Kiesler, Zdaniuk, Lundmark, & Kraut, 2000). This suggests that adults who have children whom they can ask for advice about internet use will be better skilled due to the knowledge passed on from their kids.

Studies of Web Searching

Scholars from many fields have explored how people use the Web for information retrieval. Advertising and marketing specialists usually analyze users' behavior on one particular site instead of exploring users' overall online behavior, and they focus solely on users' commercial activities online (Bell & Tang, 1998; Jarvenpaa & Todd, 1996). Researchers in the human-computer interaction field have used large-scale aggregate logs about people's Web use to analyze Web activity over a specified period (Catledge & Pitkow, 1995; Huberman, Pirolli, Pitkow, & Lukose, 1998; Spink, Jansen, Wolfram, & Saracevic, 2002). In cases where data are derived from larger segments of the online population (e.g., Hoelscher, 1998; Huberman et al., 1998; Jansen, Spink, & Saracevic, 2000; Silverstein, Henzinger, Marais, & Moricz, 1999), there is no information about specific users; thus, it is impossible to make any claims about how attributes of users may be related to their online behavior. Although some have tried to develop more general models from data on what sites people visit (Goldfarb, 2002; Sinai & Waldfogel, 2001), these studies are based on assumptions about users' behaviors that cannot be verified. They offer interesting information on users' aggregate search patterns (Spink et al., 2002), but search activities are discussed in isolation of people's social attributes.

Some researchers in the library and information science community have explored how people search for information on the Web. However, these projects limit their scope to people in particular academic communities, such as graduate students in information science programs (Wang, Hawk, & Tenopir, 2000) or, at best, college students in general (Cothey, 2002; Dennis, Bruza, & McArthur, 2002; Kim & Allen, 2002), making their findings impossible to generalize to the broader internet user population. To gain a better understanding of how the general population is using the internet, it is important to include people from beyond the academic community in such studies. Moreover, the just-cited studies rarely connect questions of Web usability with social factors; rather, they are interested in the technical specifics of search queries and users' experiences with online interfaces and tools (Spink, 2002).

These studies provide important information for a baseline understanding of how certain people navigate particular parts of the Web. However, they either limit their scope to specific user populations (e.g., information technology professionals, college students), do not collect background information about user attributes, or look at use patterns on an aggregate level without collecting data about the specific goals of a Web session. Moreover, these projects consider only people's use of search engines, ignoring the fact that the use of a search engine is only one of many ways in which to find content on the Web. In contrast, the project on which this chapter draws collects information about both user attributes and online actions. The next section describes the methodology in detail.

Methodology

The study is based on in-person observations and interviews with a random sample of internet users.[2] Respondents were given the choice of using a PC or a Mac, both of which had the three most popular browsing software applications (Internet Explorer, Netscape Communicator, and America Online) to allow respondents to replicate their usual online experience.[3] The computers connected to the internet on a high-speed university network line. In addition, a program called Don't Panic from Panicware was used to erase the browser and URL histories on each browser program so that each respondent started out with a clean slate and was not influenced by previous users' actions. The search sessions were recorded with a screen capture program that generated audio-visual files of the entire search session.[4]

Participants were given a list of tasks to perform online to see how they would find various types of information on the Web. The particular tasks were chosen to explore people's ability to find information on the Web in various topical domains. They explore whether users can find locally relevant content (specifically, information about local cultural events), the multimedia nature of the Web (e.g., music that users can listen to online), and whether users can use

the Web for political purposes (e.g., finding a site that compares various presidential candidates' views), for government information (e.g., locating tax forms), and for less common tasks (e.g., finding children's art online).

The researcher sat behind and to the left of respondents and refrained from influencing their strategies (e.g., never suggesting any particular online actions, not answering questions about spelling or whether a certain click would be useful). Respondents were encouraged to look for the information until they found it. No one was cut off from pursuing a search. In some cases where respondents looked frustrated or agitated, they were given the option of moving on. However, when respondents simply stated that they were unable to perform a certain task, they were encouraged to try several times before moving on to the next task. When respondents suggested multiple times that they would not be able to complete a search, they were presented with the next task.

Information about respondents' usual internet use and histories, as well as data on their demographic backgrounds and social support networks, was collected via surveys, one that was orally administered at the beginning of the study session and one that was filled out online at the end of the study session. Administering the questionnaire right before the observation session proved to be very useful. Because the questions explore many facets of Web use, respondents were prompted to think about numerous details of their Web experiences before sitting down at the computer and embarking on the tasks presented by the researcher.

Data

The findings reported here are based on 66 interviews conducted during the summer and fall of 2001 in the suburban towns and boroughs of a New Jersey county.[5] Respondents ranged in age from 18 to 81 years (for details, see Table 16.1). Just under half (45%) of the sample was male. More than half (58%) of the respondents worked full-time, and an additional 14% worked part-time. Their occupations included real estate agents, environmental policy analysts, blue-collar workers, office assistants, teachers, service employees, and medical professionals as well as students, unemployed persons, and retired persons.

Fully 91% of the respondents were white; there were four African American respondents and one Asian American respondent, and one respondent chose the "other" category for race and self-identified as Hispanic. In addition, 29% reported living alone, half lived with their spouses, and the rest lived with

Table 16.1 Descriptive Statistics for Independent Variables

	Mean	*Standard Deviation*	*Median*	*Minimum*	*Maximum*
Age (years)	44.26	15.65	42.5	18	81
Education[a]	N/A	N/A	College	Less than high school	Ph.D.
Family income[a]	N/A	N/A	$90,000–$99,000	$17,500–$19,000	> $250,000
Number of years since first use of the internet	6.10	3.38	6.0	0	16[b]
Amount of time spent browsing the Web weekly	9.02 hours	10.51 hours	5.5 hours	8 minutes	70 hours[c]

a. Education and family income have no means because those variables were collected categorically.

b. Number of years since first internet use was top coded at 8 for the analyses.

c. Number of hours on the Web was top coded at 56 (8 hours per day) for the analyses. Excluding this top value, the mean time spent browsing the Web per week was 7.63 hours, with a standard deviation of 6.75 hours and a median of 5.0 hours.

roommates or others (in most cases, parents). More than half (55%) of the respondents had children. Of these 36 people, 20 had children currently living with them.

On average, participants in this study were more educated than the general internet user population. This suggests that findings from the study will be conservative with respect to effects of various educational levels on people's ability to use the Web effectively. The average family income of respondents was greater than the national average, although it is important to note that this New Jersey county is one of the highest income counties in the United States; thus, despite the high median income, the sample is not out of the ordinary for the local population.[6]

Regarding Web use frequency and history, the group was diverse. Participants' Web use ranged from just a few minutes per week to more than 30 hours per week. The group was similarly diverse in respondents' overall experience with the medium. One person had gone online the year of the study, and an additional 17% had been online for 2 years or less. However, many of the respondents (42%) had been users for 5 to 7 years. There were also several long-term users among the respondents, with 14% having had their first exposure to the internet more than 10 years ago.

Table 16.2 Number of Successfully Completed Tasks

Number of Tasks Completed Successfully	*Number of Respondents*	*Percentage*
1	2	3
2	9	14
3	9	14
4	19	29
5	27	40
Total	66	100

DIFFERENCES IN PEOPLE'S ABILITY TO FIND CONTENT ONLINE

I measured people's online skills in two ways. First, the binary success/failure rate showed what portion of the respondents was able to complete each of five tasks analyzed in this study. Second, the time to completion of each task was measured in seconds to show the gradual differences in how long people take to find information on the Web. The exact time spent on each task was recorded for every respondent so that information is available on both when respondents successfully completed a task and when they decided to give up on a task.

There is some variance in people's success rates in performing tasks and large variance in the amounts of time they took to complete tasks. Of the 66 respondents, 27 (41%) were able to successfully complete all tasks, and an additional 19 (29%) succeeded in locating four of the five types of information sought (for details, see Table 16.2). However, the remaining 20 people were able to successfully complete only one to three tasks. This is a considerable proportion given that people were encouraged to pursue tasks without any time constraint.

Overall, people spent anywhere from 2½ minutes to 33 minutes on the five tasks. There was a gradual increase in the amount of time people spent on all of the tasks. Four respondents spent less than a total of 5 minutes on all five tasks, making them the expert searchers. A 26-year-old woman whizzed through the tasks by using Google in an informed manner with multiple-term queries (e.g., *Al Gore views on abortion* for the political comparison task). An 18-year-old college woman also used elaborate queries in some cases (e.g., *Bush and Gore and abortion* for the same comparison question) but turned to directory listings on Yahoo for other tasks such as finding local movie listings. People who are good searchers and who rely mostly on one search engine can get quite affectionate about their favorite online service. A 42-year-old woman who was in between jobs and who got through the five tasks very quickly commented at the start of one task, "I go back to Google, my old friend."

Four respondents spent more than a total of 30 minutes on the five tasks, exhibiting very limited Web user skills. Simply using Google is not enough. One 56-year-old man who worked in business development clicked on the Advanced Search feature of Google and then typed *abortion comparison Bush and Gore* in the search field labeled "with the exact phrase"; this proved to be too nuanced and yielded no results. Confused by this, he clicked on the search button a few more times—without changing anything in the search field—and continued to get no results. After some additional attempts at using that feature of the search engine, he moved on to CNN's site and found the information via its search engine.

An action typical of those who understand little about the Web and searching is to continuously click on "related search" items after a search result. This leads to yet more search results. If the user simply keeps clicking on these links, he or she never gets to an actual site. A 44-year-old woman who worked in retail repeatedly used this method to find sites and took many iterations of searches to finally arrive at a page with content of interest. There are also those who never use search engines. These users rely on site recommendations from their internet service providers.

EXPLAINING DIFFERENCES IN PEOPLE'S ONLINE SKILLS

To look at what accounts for the differences in whether people are able to find certain types of content online, I estimated probit models of task completion.[7] I have valid completion data for 317 tasks of the total 330 tasks performed by the 66 individuals (data are missing for 13 tasks). Table 16.3 presents the estimates of the probit models normalized to represent the marginal effect at the average of the explanatory variables on the probability of completion of tasks. Model 1 includes demographic variables only. Results suggest that age was negatively associated with people's ability to complete tasks successfully. In contrast, having a graduate degree was significantly associated with people's ability to find various types of content online. Model 2 adds information about people's prior experience with the internet by including data about the amount of time people spend surfing the Web each week and the number of years since the respondents first used the internet.[8] The results suggest that the latter was positively associated with users' ability to find information on the Web.

Table 16.3 Probabilities of Successful Completion of Tasks (probit analysis)

	Model 1:		*Model 2:*	
	Demographics	*Standard Error*	*Demographics and Online Experience*	*Standard Error*
Demographics				
Age	−.006**	.002	−.005***	.002
Gender (female)	.058	.147	−.032	.051
Age × Gender interaction	−.001	.003		
Family income (logged)	.063	.041	.057	.037
Less than college education	−.007	.069	−.021	.067
Graduate degree	.126**	.052	.084	.055
Recently in school	.024	.086	.021	.080
Children in household	−.090	.082	−.101	.080
Children × Gender interaction	−.012	.101	−.040	.100
Internet experience				
Amount of time spent browsing the Web per week			−.001	.002
Number of internet use years			.022**	.010
Chi-square	0.0000		0.0000	
Log likelihood	−132.061		−129.932	
N	300		300	

**$p \leq 0.05$, ***$p \leq 0.01$.

To examine what explains differences in how long people take to complete tasks, I specified a linear regression model of the log time to completion. I used a normal-censored data model for estimation due to the fact that we cannot observe the actual time to completion for those who gave up (i.e., right-censored data).[9] Table 16.4 presents the results of the censored normal regression models.[10] First, I looked at the predictive power of the demographic variables on their own.[11] The older people were, the longer they took to complete tasks. Education was also related to people's ability to find information online quickly; those with graduate degrees were considerably faster than those without graduate degrees. The payoffs of higher education were especially relevant for those who obtained their college and graduate degrees during the years since the Web has been widely available on university campuses. However, the kind of task presented had significant effects on the time to completion. Music-related tasks were completed more quickly, whereas people needed more time to complete the political search task.

In Model 2, I added information about previous experience with the technology to control for amount of time spent online and number of years of experience with the medium. Those who spent more time online each week were quicker at finding content on the Web. Moreover, each additional year of experience with the internet led to a few seconds of improvement in searches. The findings from the previous model were robust. Having a graduate degree helped a respondent's ability to quickly find content online by resulting in searches that were, on average, 49 seconds quicker, and this payoff was especially strong for those who attended school during recent years (resulting in searches that were more than a full minute quicker). Older people were slower, and the presence of children in the household was also associated with less efficient searching behavior, even when controlling for the amount of time spent online each week. In Model 1, the interaction of children in the household with gender was significant, suggesting that the presence of children especially takes a toll on women's online performance. However, after controlling for time online each week, the interaction term lost significance.

The presence of children in the household added an average of 39 seconds to each search task. This effect is the opposite of the predicted outcome; the expected relationship suggested that the presence of children would be beneficial to people's internet skills due to knowledge passed on from kids to parents. Rather, it seems that parents may simply hand over tasks to their offspring when in need of help. Kraut, Scherlis, Mukophadhyay, Manning, and Kiesler (1996) found that in households with children, teenagers tend to dominate use of the network-connected machines. The data presented in this chapter suggest that young users tend to be much more skilled than older users. Because of the presence of other experts in the household, parents might not develop certain online skills, preferring to delegate the tasks of looking up certain information to their children.

Table 16.4 Linear Regression Model of the Log Time to Completion of Tasks (standard errors in parentheses)

	Model 1:		*Model 2:*	
	Demographics	*Standard Error*	*Demographics and Online Experience*	*Standard Error*
Constant	4.823***	.478	5.155***	.484
Demographics				
Age	.011*	.006	0.012**	.005
Gender (female)	−.248	.323	.080	.123
Age × Gender interaction	.006	.007		
Less than college education	−.227	.143	−.195	.142
Graduate degree	−.529***	.120	−.395***	.127
Recently in school	−.469***	.179	−.410**	.180
Children in household	.307*	.161	.329**	.159
Children × Gender interaction	.364*	.217	.217	.220
Internet experience				
Amount of time spent browsing the Web per week			−.003	.005
Number of internet use years			−.073***	.025
Tasks				
Music task	−.480***	.153	−.486***	.151
Political task	1.110***	.189	1.120***	.157
Tax task	.144	.150	.144	.149
Children's art task	.110	.150	.113	.148
Uncensored	242		242	
Right censored	56		56	
Pseudo-R^2	0.1951		0.2058	
N	298		298	

NOTE: The task variables are dichotomous.

*$p \leq 0.10$, **$p \leq 0.05$, ***$p \leq 0.01$.

This finding suggests the importance of a more in-depth look at the influence of people's social relationships on online behavior. Here, I turn to the qualitative data derived from the conversations with participants during the observation sessions to tease out the role of social ties and the influence of people's social surroundings on their Web use. As suggested earlier, child rearing may cut into people's Web use by limiting the amount of time people have to explore the Web freely (Kelsey, 2002; Lally, 2002). A 39-year-old, stay-at-home mother successfully completed the five tasks analyzed in this study but was one of the slowest to do so (she took 25 minutes, meaning that she was slower than 90% of the respondents). She lived with her husband and three children and noted the following about how her family situation may influence her online behavior:

> I think having children definitely affects my Web use in that I just don't have the time to sit at the computer and spend the time that is takes to figure out how to get around. It is extremely frustrating to try and concentrate with kids around.

Another related and seemingly important factor is how new media resources are allocated within the household. Some parents commented on how their children's time spent online influences their own ability to access the Web. A 48-year-old widow, a social worker who lived with her 15-year-old daughter, described her access to the computer in the context of her daughter's internet use: "My internet usage is affected by the amount of time [my daughter] is on the internet. I have to ask her to get off so I can just check my e-mail a couple of times each week." Although this participant successfully completed all of the tasks, she was among the slowest 25%.

In contrast, a 39-year-old man, the information systems director of a small business, has created an environment at home that allows all members of his family full access to the Web at their leisure:

> We have four computers in the house: one PC for my wife and I, one PC for my 16-year-old, one PC for my 8-year-old, and a fourth PC on the network with 2 LPT ports on it that acts as a print server; we have a laser printer and a color ink jet. It also acts as a shared drive for exchanging files on the network. It also has access to the internet, so even if my wife and both kids are on the Web, I can do a quick search or surf for whatever I need on the print server PC.

In addition to these resources at home, this participant also said he had access to two computers at work. It is not surprising that he succeeded with all five tasks and finished among the top 25% of users in speed.

Admittedly, few have access to such an abundance of technical resources. Instead, others rely on their social ties to improve their online techniques. They turn to family and friends for site recommendations and specific advice with particular queries. A 47-year-old public school teacher who succeeded with four of the five tasks and placed in the slower half of the group explicitly stated that her children are the "primary source of my knowledge. And my husband, too. They certainly know more than I do." In the study setting, her family members were not with her, so she had to rely on memory of what more general strategies she may have picked up from them.

A 39-year-old man who assembles air conditioners for a living relied on public transportation to come to the study location. Several times when he was unable to perform a task, he referred to his nephews:

> I don't know how to do that. I've seen my nephews and them do it and my daughter. They on the computer all the time. . . . I even have a nephew that's, boy, 6 years old. . . . He will get up out of bed and get on the computer.

This man attempted online shopping once and completely relied on his nephew's help to place a book order. However, the order was not successful; the book never arrived despite the fact that he had a money order made out for it. This participant was unable to complete three of the five tasks in the absence of his support network.

A younger respondent, a 20-year-old college student who successfully completed all five tasks and was the second quickest of all respondents, was asked about her use of a particularly powerful search engine. She noted, "One of my friends told me about it, and so I've been using it a lot." A 40-year-old professor who quickly found the Web site of a radio station for music commented that he had first heard about it from a friend.

From the preceding quotes, it seems that people's Web use is significantly influenced by their social surroundings. Users turn to their social support networks for advice on use of the internet. Although constrained by how much autonomy they have otherwise—the available time to explore the medium and access to a machine whenever convenient—they rely on recommendations and advice from their family and friends for more efficient use of the Web. Consequently, those who do not have such support networks to draw on miss out on important ways in which they could improve their online abilities.

Conclusion

In general, young people (those in their late teens or 20s) have a much easier time in getting around online than do their older counterparts (whether in their 40s or their 70s). Education is also significantly associated with online skills. Some of these relationships are clearly based on comfort with the technology and are not necessarily based on elaborate techniques that people have mastered specifically with respect to the Web. Given that the amount of time people have been internet users also affects how well they are able to navigate the contents of the Web, it is possible that those who are currently less skilled will learn over time and improve their ability to find content online. However, people may be discouraged by the difficulties of finding information on the Web and so may end up spending less time with the medium, resulting in continued lower level skills. Moreover, those who can draw on family and friends for advice have continuing support to help improve their online skills or, at least, to alleviate some of the frustrations caused by confused Web searching.

Using a mix of quantitative and qualitative methods to analyze the online actions of a random sample of internet users allows us to uncover generalizable trends in people's online use while also understanding the more intricate reasons for their particular online actions. People rely on their social support networks for suggestions of site recommendations and for answers to particular questions that come up as they look to the Web for answers to questions or for mere entertainment. By using a mix of methods for data collection, it is possible to give voice to users without having to make unfounded assumptions about their online behavior based on aggregate data. The quasi-experimental methods used in this study can be replicated to look at other aspects of people's use of the internet and other new communication technologies, resulting in rich data allowing nuanced analyses of processes underlying how these new media are being incorporated into people's everyday lives.

Notes

1. Because of the small number of minority participants in this study, it is not possible to explore the relationship between race/ethnicity and people's online skills based on the data in this study.

2. This approach has both advantages and shortcomings. Requesting users to come to a location will affect response rates. It also places people in an unfamiliar location and requires them to use a computer that is configured differently from the machines they usually use for browsing. This may influence the results because certain settings (e.g., the default homepage, bookmarks) are

not equivalent to their own. However, this approach controls for quality of internet connection and hardware/software differences. It also allows the researcher to concentrate on Web use knowledge in a setting that is equally different and new for all. Moreover, using one computer allows for the setup of particular software applications that are required for data recording. To be able to generalize from the findings, respondents were recruited through a random sample of residential addresses obtained for a suburban county in New Jersey from Survey Sampling Inc. This list was then checked against the National Change of Address Database maintained by the U.S. Postal Service to minimize the number of nonresidential and out-of-service numbers. Potential respondents were first mailed a letter and brochure explaining the project and requesting participation. They were directed to the study's Web site (www.webuse.org) for more information and were given the option of calling or writing the researcher to schedule an appointment. A few days after the letters had been sent, the households were contacted by telephone. In each household, the eligible English-speaking internet adult over 18 years of age with the next nearest birthday was selected to randomly sample from within households. If this person was not willing to participate, the household was coded as a refusal even if another member of the household would have been willing to take part in the study. Such strict standards help to achieve a representative sample of the internet user population. Web users were defined as people who go online at least once every month for more than just using e-mail. This is a low threshold for including people in the study and is used to maximize variance in experience (for the distribution of participants' weekly Web use, see the final row in Table 16.1). For their participation, respondents were offered $40, which they received after the observation session. Respondents were asked to come to the research site on the university campus and were offered assistance with transportation if they could not provide their own. One person took advantage of this option and was reimbursed for bus fare. The response rate was 48%, considerably high given the time cost involved in participation.

3. No default page was set on browsers so as not to influence respondents' initial actions once online. The sessions were started off by the researcher asking respondents to recall, if possible, the default homepages on the computers they use most often. The respondents were also asked whether they had personalized anything on their browsers and whether they had set any bookmarks or favorites.

4. The Hypercam program from Hyperionics was used on the PC, and the SnapZPro e program from Ambrosia Software was used on the Mac. The whole screen was captured, as was every action (e.g., click of the mouse, scrolling) and every verbal comment respondents made during their searches. These files were then analyzed to measure whether people successfully completed a task and how long they took to do so.

5. I conducted 55 of the interviews, and a research assistant conducted 11.

6. The average median per capita income in this county was nearly $40,000 in 2000 (based on census data). Here, I look at household income, which is likely to be considerably higher on average. Moreover, because this sample excludes the inner-city population of the county's biggest town due to logistical purposes of the study, the poorest neighborhoods are not in the sample, suggesting that the median per capita income in the county is considerably higher than $40,000.

7. I used the "cluster" command in the Stata statistical software package due to repeated observations on individuals. Without clustering, error terms would not be independent across tasks due to individual characteristics of respondents not accounted for in the model.

8. I removed the interaction of age and gender from the second model because it did not prove to be significant in the initial model.

9. This model was estimated by maximum likelihood using the command "cnreg" in the Stata statistical software program.

10. A dummy variable was included in four of the five tasks to control for task-specific effects.

11. It is important to note that times-to-task within one person will be correlated based on the person's individual characteristics not accounted for in the model. This results in error terms that are not independent across tasks within a person.

References

Bell, H., & Tang, N. K. H. (1998). The effectiveness of commercial internet Web sites: A user's perspective. *Internet Research, 8,* 219–228.

Catledge, L. D., & Pitkow, J. E. (1995). *Characterizing browsing strategies in the World Wide Web.* Paper presented at the Third International World Wide Web Conference, Darmstadt, Germany.

Cothey, V. (2002). A longitudinal study of World Wide Web users' information-searching behavior. *Journal of the American Society for Information Science and Technology, 53*(2), 67–78.

Dennis, S., Bruza, P., & McArthur, R. (2002). Web searching: A process-oriented experimental study of three interactive search paradigms. *Journal of the American Society for Information Science and Technology, 53*(2), 120–133.

DiMaggio, P., Hargittai, E., Neuman, R., & Robinson, J. (2001). Social implications of the internet. *Annual Review of Sociology, 27,* 307–336.

Dutton, W. H., Rogers, E. M., & Jun, S-H. (1987). Diffusion and social impacts of personal computers. *Communication Research, 14,* 219–250.

Goldfarb, A. (2002). Analyzing Website choice using clickstream data. In M. R. Baye (Ed.), *Advances in applied microeconomics,* Vol. 11: *The economics of the internet and e-commerce* (pp. 209–230). Amsterdam: Elsevier Science.

Hoelscher, C. (1998). *How internet experts search for information on the Web.* Paper presented at the Conference of the World Wide Web, Internet, and Intranet, Orlando, FL.

Howard, P. E. N., Rainie, L., & Jones, S. (2001). Days and nights on the internet: The impact of a diffusing technology. *American Behavioral Scientist, 45,* 383–404.

Huberman, B. A., Pirolli, P. L., Pitkow, J. E., & Lukose, R. M. (1998). Strong regularities in World Wide Web surfing. *Science, 280,* 94–97.

Jansen, B. J., & Pooch, U. (2001). A review of Web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology, 52,* 235–246.

Jansen, B. J., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: A study and analysis of user queries on the Web. *Information Processing and Management, 36,* 207–227.

Jarvenpaa, S. L., & Todd, P. A. (1996). Consumer reactions to electronic shopping on the World Wide Web. *International Journal of Electronic Commerce, 1*(2), 59–88.

Kelsey, D. (2002, January 18). U.S. women's net use grows at triple the rate of men's. *NewsBytes.* Retrieved May 10, 2003, from www.washingtonpost.com

Kiesler, S., Zdaniuk, B., Lundmark, V., & Kraut, R. (2000). Troubles with the internet: The dynamics of help at home. *Human-Computer Interaction, 15,* 323–351.

Kim, K-S., & Allen, B. (2002). Cognitive and task influences on Web searching behavior. *Journal of the American Society for Information Science and Technology, 53*(2), 109–119.

Kraut, R., Scherlis, W., Mukophadhyay, T., Manning, J., & Kiesler, S. (1996). The HomeNet Field Trial of Residential Internet Services. *Communications of the ACM, 39*(12), 55–63.

Lally, E. (2002). *At home with computers.* Oxford, UK: Berg.

Loges, W. E., & Jung, J-Y. (2001). Exploring the digital divide: internet connectedness and age. *Communication Research, 28,* 536–562.

Murdock, G., Hartmann, P. & Gray, P. (1992). Contextualizing home computing: Resources and practices. In R. Silverstone & E. Hirsch (Eds.), *Consuming technologies* (pp. 146–160). New York: Routledge.

National Telecommunications and Information Administration. (1995). *Falling through the net: A survey of the "have-nots" in rural and urban America.* Washington, DC: Author.

National Telecommunications and Information Administration. (1998). *Falling through the net II: New data on the digital divide.* Washington, DC: Author.

National Telecommunications and Information Administration. (1999). *Falling through the net: Defining the digital divide.* Washington, DC: Author.

National Telecommunications and Information Administration. (2000). *Falling through the net: Toward digital inclusion.* Washington, DC: Author.

National Telecommunications and Information Administration. (2002). *A nation online.* Washington, DC: Author.

Rogers, E. M. (1985). The diffusion of home computers among households in Silicon Valley. *Marriage and Family Review, 8,* 89–101.

Rogers, E. (1995). *Diffusion of innovations.* New York: Free Press.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web search engine query log. *SIGIR Forum, 33*(1), 6–12.

Sinai, T., & Waldfogel, J. (2001). *Geography and the internet: Is the internet a substitute or a complement for cities?* Paper presented at the 29th Annual Telecommunications Policy Research Conference, Alexandria, VA.

Spink, A. (2002). Introduction to the special issue on Web research. *Journal of the American Society for Information Science and Technology, 53*(2), 65–66.

Spink, A., Jansen, B. J., Wolfram, D., & Saracevic, T. (2002). From e-sex to e-commerce: Web search changes. *Computer 35*(3), 107–109.

Wang, P., Hawk, W. B., & Tenopir, C. (2000). Users' interactions with World Wide Web resources: An exploratory study using a holistic approach. *Information Processing and Management, 36,* 229–251.

17

American Internet Users and Privacy

A Safe Harbor of Their Own?

Doreen Starke-Meyerring
University of Minnesota

Dan L. Burk
University of Minnesota

Laura J. Gurak
University of Minnesota

The rise of popular internet use has been accompanied by a parallel rise in concern over personal and data privacy. According to a 1999 *Wall Street Journal* poll, for example, privacy violations rank among the top concern of Americans. Similarly, a Lou Harris poll found that the percentage of respondents concerned about their privacy rights surged from 34% in 1970 to 90% in 1998 (Identity Theft Resource Center, 2002). The Pew Internet and American Life Project (2000) report, *Trust and Privacy Online: Why Americans Want to Rewrite the Rules,* underscored this new intensity of privacy concerns. According to the report, 84% of respondents would be "very concerned" or "somewhat concerned" about their personal information possibly falling into the wrong hands (p. 22).

The Pew report also provided valuable insights into the online privacy practices and wishes of American internet users such as their expectations about the privacy of their personal information, their knowledge about privacy or

For information:

Sage Publications, Inc.
2455 Teller Road
Thousand Oaks, California 91320
E-mail: order@sagepub.com

Sage Publications Ltd.
6 Bonhill Street
London EC2A 4PU
United Kingdom

Sage Publications India Pvt. Ltd.
B-42, Panchsheel Enclave
Post Box 4109
New Delhi 110 017 India

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Society online : the internet in context / edited by Philip N. Howard and Steve Jones.
p. cm.
Includes bibliographical references and index.
ISBN 0-7619-2707-7 (Cloth) — ISBN 0-7619-2708-5 (Paper)
1. Information society. 2. Internet. I. Howard, Philip N. II. Jones, Steve, 1961-
HM851.S655 2004
303.48'33—dc21

2003008611

Printed on acid-free paper

03 04 05 06 07 08 09 10 9 8 7 6 5 4 3 2 1

Acquiring Editor: Margaret H. Seawell
Editorial Assistant: Alicia Carter
Production Editor: Claudia A. Hoffman
Typesetter: C&M Digitals (P) Ltd.
Copy Editor: D. J. Peck
Indexer: Molly Hall
Cover Designer: Edgar Abarca

Contents