

SERVING CITIZENS' NEEDS: MINIMIZING ONLINE HURDLES TO ACCESSING GOVERNMENT INFORMATION

ESZTER HARGITTAI

ABSTRACT

With the rapid spread of the Internet across society, government institutions are taking advantage of digital technology to distribute materials to citizens. Is merely having a Web site enough, or are there certain usability considerations site creators must keep in mind to assure efficient public access to online materials? This project looked at typical people's ability to locate various types of content online, in particular, their ability to find tax forms on the Web. Findings suggest that people look for content in a myriad of ways, and there is considerable variance in how long people take to complete this online task. Users are often confused by the ways in which content is presented to them. In this paper, two common sources of confusion in users' online experiences with locating tax forms online are distinguished: (1) URL confusion and (2) page design layout. Ways are also suggested to decrease these two sources of frustration, yielding less exasperating and more productive user experiences.

Eszter Hargittai is a graduate student in the Sociology Department, Wallace Hall, Princeton University, Princeton, NJ 08544. eszter@phoenix.princeton.edu.

Special thanks to: Paul DiMaggio for his insightful comments throughout this project, Stan Katz for his ongoing support, Erica Field and John Robinson for helpful discussions, Susan Lutz and Inna Barmash for their assistance with data collection, Hank Farber and Betty Leydon for their logistical help, and the many people who took time from their busy schedules to participate in this study. Generous support from the Markle Foundation and NSF grants #SES9819907 and #ITR0086143 is kindly acknowledged. The project has also been supported in part by a grant from the Russell Sage Foundation and through a grant from the Pew Charitable Trusts to the Center for Arts and Cultural Policy Studies, Princeton University. The author is also grateful to the Dan David Foundation for its support.

There are many studies that look at how people use the Internet (DiMaggio et al. 2001) and in particular what types of content they view online (Howard, Rainie and 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 (Jansen and 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. Moreover, such literature rarely considers the production side of online content distribution. Does the way in which content is organized and presented online influence typical people's ability to find their way to material on the Web? This article reports findings from a project that explores people's search strategies in locating content on the Web, in conjunction with a look at how Web sites present information to users. The focus is on access to government documents, in particular, how users find individual federal tax forms on the Web.

The first section links this topic to the literature on digital inequality. The next section outlines the following: the methods used for studying how people locate content online, the sampling methodology, and the method for analyzing the data. Findings are reported from 97 interviews and in-person observations conducted with a random sample of Internet users. A particular focus is on how URL confusion and page design layout problems contribute to frustrating experiences for users. Finally, ways are suggested in which these two sources of confusion can be remedied to offer a more productive and satisfying user experience—not only for locating tax forms online, but also for locating other types of government documents or any other type of material on the Web.

DIGITAL INEQUALITY

Millions of people access the Web daily for financial, health and government information, for job searches, for entertainment and for pursuing numerous other activities (Howard, Rainie and Jones 2001). As daily activities continue to move online, the ability to efficiently navigate the Web for information becomes increasingly important to maintaining a competitive edge and guaranteeing equal opportunity in the various aspects of life—including access to government, health and financial services, education, consumer information and a myriad of other realms. As such, Web-use skills are poised to become an important part of people's human capital. Internet skills are a new, vital component of human capital in the digital age.

Much attention among both academic researchers (Bucy 2000; Hoffman and Novak 1998; Strover 1999) and in policy circles (Benton Foundation and Leadership Conference on Civil Rights Education Fund 2002; The National Telecommunications and Information Administration 2002) has been paid to what segments of the population have access to the Internet or are Internet users. Access is usually defined as having a network-connected machine in

one's home or workplace. "Use" -more specifically refers to how people actually make use of the medium, after gaining access to it. The *digital divide* is most often conceptualized in binary terms: someone either has access to the medium or does not; someone either uses the Internet or does not. Such an approach to the *digital divide* has led some to conclude that we can "declare the war won," given that access and use have increased consistently over the past years (NTIA 2002) and the majority of the American population is now online (Compaine 2001).

However, this approach wrongly assumes that gaining access to the Internet obliterates any potential inequality that may result from lack of access to the new medium. There are factors beyond mere connectivity that need to be considered when discussing the potential implications of the Internet for inequality. In addition to relying on basic measures of access to a medium, one needs to consider more nuanced measures of use such as user "skill". "Skill" is defined as the ability to locate content online effectively and efficiently.

In addition to refining our understanding of what it means to have effective access to a medium (Wilson 2000), it is also important to consider the institutional factors that influence how people use a medium. Individual uses are embedded in a higher-level system organization. Business practices shape the landscape of a new technology's industrial organization, which then affects how it is made available, presented and distributed to users. This research report discusses findings about online information-seeking, both at the user-skill level and with respect to organizational practices that influence users' online actions.

STUDY DESIGN AND METHODOLOGY

The study is based on in-person observations and interviews with a random sample of Internet users from a New Jersey county. Respondents were asked to come to a university location for participation, where they were given the choice of using a PC or a Mac. Both computers were loaded with the three, most popular browsing software applications (Internet Explorer, America Online, and Netscape Communicator), allowing respondents to replicate their usual online experience.¹ The computers were connected to the Internet on a high-speed, university network line. Additionally, a program called "Don't Panic" (Panicware 2001) was used to erase the browser and URL history 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.²

The researcher sat behind to the left of the respondent and refrained from influencing the respondents' strategies (e.g., never suggesting any particular online action, not answering questions about spelling or saying 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 cases where respondents looked frustrated or agitated, they were given the option to move on to another task. However, subjects who simply stated that they were unable to perform a certain task were encouraged to try several times before moving on to the next task.

Information about subjects' usual Internet use and history and data on demographic background were collected via surveys, one orally administered at the beginning of the study session and the other self-completed online at the end of the study session. Audio-visual files were coded to see whether people successfully completed tasks and the number of seconds they took to do so. Information about what sites users visited and which online steps led them there was also coded. The audio component of the interviews was transcribed to offer information about users' understanding of their actions and about their feelings about their search experience. Hargittai (2002) describes the complete study methodology in more detail.

The findings reported here are based on 97 observation sessions and interviews conducted between the summers of 2001 and 2002 in a New Jersey county.³ Unlike respondents in similar studies (McDonald and Spencer 2000; Wang, Hawk and Tenopir 2000), the participants in this study represent a diverse group of Internet users. As shown in Table 1, participants range in age from 18–81, about half (52%) being women (see Table 1 for details). Participants' occupations range from real-estate agent, environmental policy analyst, blue-collar worker to office assistant, teacher, service employee and medical professional, student, unemployed and retired.

The group is also diverse regarding Web use frequency and history. Participants' Web use ranges from just a few minutes a week to over 30 hours weekly. The group is similarly diverse in its overall experience with the Web medium. One person first went online the year of the study, and an additional 13 percent had used the Internet for two years or less. However, almost 40% of the subjects had been users for 5–7 years, and 15% of respondents had first been exposed to the Internet more than ten years ago.

Differences in search effectiveness and efficiency: Participants were asked to look for tax forms online, in particular, for the individual, federal 1040 form. Respondents used a diverse set of methods to find such government material. When given unlimited amount of time, 93 percent of participants were able to locate such a tax form on the Web. The seven participants who were unable to find tax forms online tended to be statistically significantly older people who spent considerably less time online per week and had started to use the Internet more recently than those who could successfully complete this task.

There is a large variance in the amount of time respondents spent on this task. Users took anywhere from less than half a minute to almost nine minutes to look for tax forms on the Web, with the average participant spending two and a half minutes searching for such a form. The analyses controlled for earlier

TABLE 1: BACKGROUND INFORMATION ABOUT RESPONDENT SAMPLE

	Mean	St. Dev.	Median	Minimum	Maximum
Age	43	15.9	42	18	81
Education*	N/A*	N/A*	College	Less than high school	Ph.D.
Family income [†]	N/A*	N/A*	\$80-89,000 [†]	\$17,500-19,000	>\$250,000
Years using the Internet	6.3 yrs	3.4 yrs	6 yrs	0 yrs	16 yrs
Hours of weekly Web browsing	8.6 hrs	9.4 hrs	7 hrs	0.1 hrs	70 hours
* Education and Family income have no means as those variables were collected categorically.					
[†] Average median household income in this county in 2000 was almost \$60,000 (based on Census data) with modal household income of \$75,000–99,999; thus this sample is what one would expect for Internet users in this area.					

experience with looking for tax information online, but having such past experience had no significant effect on successfully completing the task.

SEARCHING FOR TAX FORMS ONLINE

Types of search strategies: Users turned to a variety of strategies to begin their quest for the form, with 60% of users turning to a search engine and 40% to a specific URL (*Uniform Resource Locator* or Web address). Among those who used a search engine, 29% turned to Google, 24% to AOL, 16% to MSN's search engine, 15% to Yahoo! and the rest to various other search engines and portal sites. The majority of those who turned to Google or Yahoo for their search tended to be quicker in completing this task, whereas those who relied on AOL or who went to a topic specific site (as opposed to a search engine or portal site) as their first move tended to take longer in successfully finding tax forms online.

Among those who tried a specific URL, the majority tried to guess the address of the Web site for the Internal Revenue Service. The guesses included "irs" with the following three extensions: .gov, .com, and .org. The correct address is www.irs.gov, yet only 52 percent of those trying a URL used that address. It turns out that using this strategy did not improve one's chances of finding tax forms due to design concerns on the IRS's Web site (these are discussed in the next section). However, with better usability on the IRS's Web site, knowing to go directly to www.irs.gov could aid in people finding tax forms quicker than those who didn't know the correct online address of the agency. As shown in Table 2, 33% of participants tried to go to www.irs.com, and 15% entered www.irs.org in the location bar of their browser. In 2001, the .com and

TABLE 2: SEARCH STRATEGIES USED WHEN LOOKING FOR TAX FORMS ONLINE

Strategy	Percentage of users	IRS domain extension	Percentage of users
IRS-related URL	30	.gov	52
Search engine	60	.com	33
Specific site (other than "irs")	10	.org	15

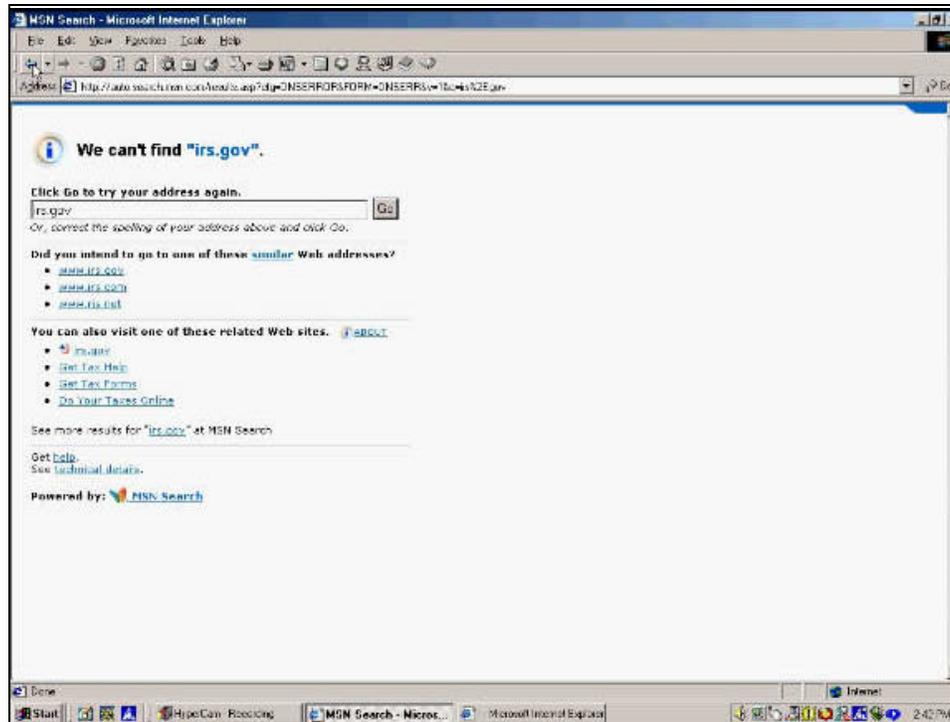
.org addresses redirected the user to the same commercial site, but in 2002, they led to two different sites, neither affiliated with the Internal Revenue Service. Some people tried yet other variations of irs, e.g., "irs.dot.gov" or simply "www.irs" without any extensions for specifying the top-level domain. Among participants trying to access a site directly, a handful went to a financial site with which they were previously familiar. These people soon switched to search engines, because they could not find tax forms on these sites.

SOURCES OF CONFUSION AND FRUSTRATION

Many users expressed considerable frustration with this task, despite the fact that tax forms are available online in a myriad of locations and that the Internal Revenue Service itself has a comprehensive site with many forms and other information. There were two major sources of frustration for users: (1) confusion stemming from the use of various Web addresses (anything other than "www.irs.gov") and (2) misunderstandings or uncertainty caused by the organization and layout of the IRS's Web site. These two points are discussed in detail below.

URL confusion: As mentioned above, over half of the users knew the correct address of the IRS's Web site: www.irs.gov. However, note that three of the people who knew to go to "irs.gov" failed to enter "www" as the beginning of the address. To this day, one encounters an error when attempting to access the IRS without the "www" in the address, as shown in the Figure 1 screen shot of what happens when "irs.gov" is entered in Internet Explorer.⁴

FIGURE 1: SCREENSHOT OF ERROR AFTER ENTERING “IRS.GOV” IN THE LOCATION BAR OF INTERNET EXPLORER (FALL, 2001)



A more common source of confusion for users was the busy, commercial site that comes up under “www.irs.com”. Until recently, www.irs.org also redirected to this page. The Figure 2 screen shot for this site features numerous links, but none of them links directly to actual forms. Although there is a link called “Individual Tax Forms,” it leads to a search engine’s “results” page containing several, more confusing commercial links as shown in Figure 3.

A 36-year-old woman who works for a local government agency during her search process commented:

[..] Oh ! dot com, no dot gov...wow look at that...that’s fancy...so... “get your free”...what is this? [laughs] this is so silly! [laughs]...I mean is this stupid or what? Oh here “department of the treasury”, let me just...oh my gosh! I don’t know! [...] Ok, wait...oh my gosh...here... www.irs.gov...ok, so you get to this ridiculous [laughs] it’s not funny [laughs] “tax stats, tax”...oh maybe “electronic services”?...well this is to find, to do it online? Oh just to find forms? So watch, so I clicked on that...oh ok...that didn’t come up though... “enter another government”, “you will leave the”...oh! “enter another government Web site created, operated, and maintained by that agency”... “continue”...[15 second pause] well that’s not very helpful...ok, so I would probably go to [...] I

FIGURE 2: SCREENSHOT OF WWW.IRS.COM (FALL, 2001)

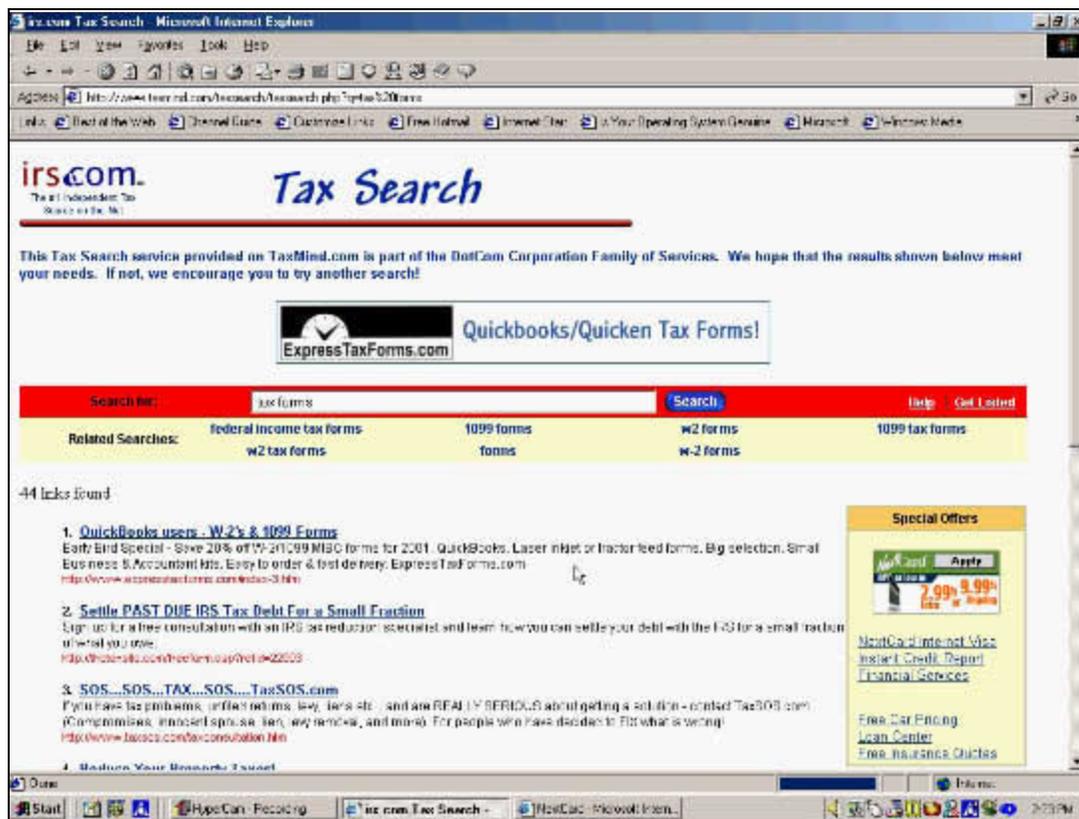


would probably just go back to a search engine and type in...[laughs]... “free 1040 tax forms”...free? [excited] Of course they’re free...you have to pay for tax forms? C’mon...I get all excited when I see the word “free”, and then I’m like “what the heck?” [laughing] hello!... “1040 easy”, “tax return online”... “quick tax” see I guess I would search for something that had “dot gov” behind it... “file manager for forms” maybe... “forms and publications” but is that...see that’s...oh! Ok that only took about 20 minutes! [laughs]

This entire search took about 3 minutes, but the respondent’s comment at the end suggests that she perceived the search process to have taken longer than expected, given the task. As she narrates most of what she does, one can see that she is quite confused about many of the steps in the process. The various commercial sites that came up as she entered www.irs.org did not lead her to the forms for which she was looking. Several other participants encountered the same problem of URL confusion and spent several minutes clicking from one irrelevant site to another.

Web site organization and design problems: The effect of design on users is well-captured by one user’s reaction to the IRS’s 2001 site: “I would go to www.irs.gov [she types it into the location bar and the IRS’s site comes up], I believe that is it. *And that is not it.*” [author’s emphasis]. This user, a 25-year-

FIGURE 3: SCREENSHOT OF SITE LINKED FROM WWW.IRS.COM (SUMMER, 2001)



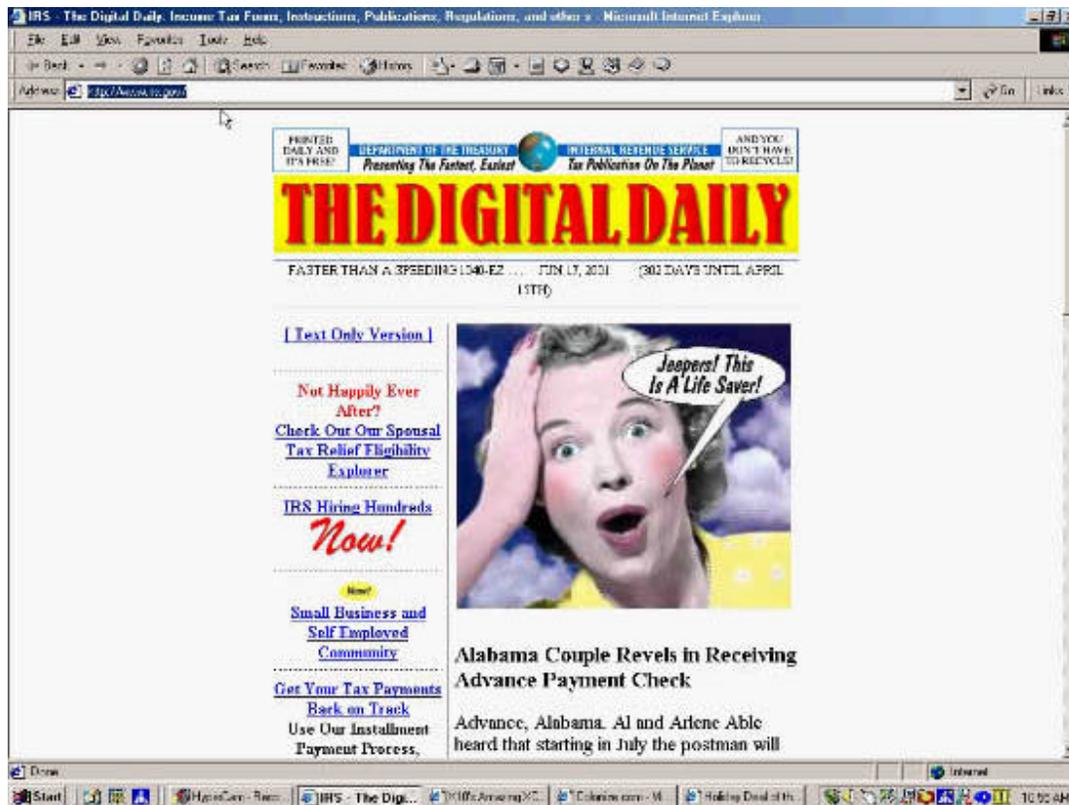
old college-educated woman who works as a financial analyst, was so confused by the design on the IRS's welcome page (See Figure 4) that she immediately hit the Back button on her browser and used Yahoo's search engine instead.

Another user, a 40-year-old man who works on medical equipment repair and installation, had a similar impression: "It looks almost [...] comical [...] to me it doesn't look like a professional site, I don't know...[laughing] I wouldn't go there to file income taxes or to get a tax form...[laughing] would you?"

Those users who did recognize that they were on the IRS's site were confused at later stages in their search. The quickest path to tax forms on the 2001 version of the IRS site was to click on Forms and Publications and then follow a link called Forms and Instructions. Unfortunately, the "Forms and Pubs" link was available only at the bottom of the IRS's home page—well below what a user would be able to see when the page initially came up in her browser. Not everyone saw this link, and so users clicked something else that could possibly lead to forms on the Web site, but this involved a guessing game.

Such ambiguity is frustrating for users. A 69-year-old, retired marketer of scientific equipment commented with irony: "This is crazy; your government

FIGURE 4: SCREENSHOT OF WWW.IRS.GOV (SUMMER, 2001)

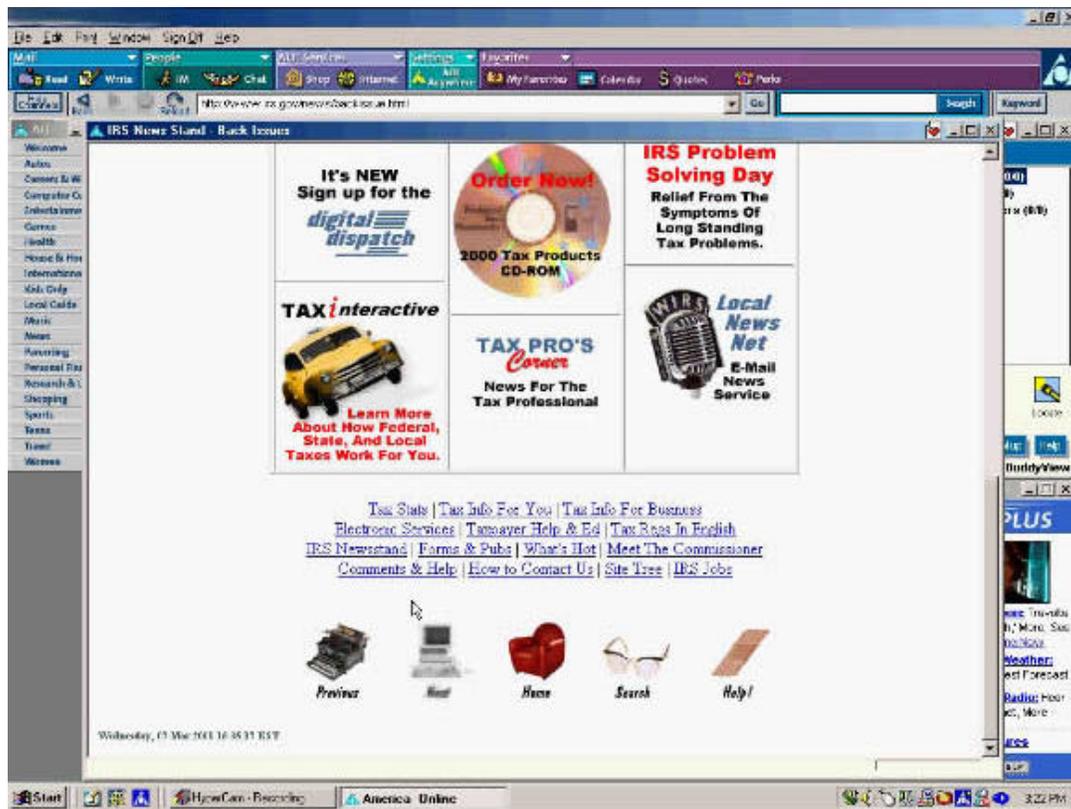


working for you” after frustratingly scrolling up and down on a page that seemed like it might have forms but didn’t (see Figure 5). In the end, he did notice the “Forms and Pubs” link on the bottom of the page, but only after getting somewhat frustrated with the overall experience.

Even those who quickly found their way to the Forms and Instructions page were in for the ultimate surprise, in that there was no obvious, visible link to forms themselves. Figure 6 shows the part of the Forms and Publications page on the 2001 IRS site that was viewable on a standard 17” monitor upon entry to the page. The entire section explains only how one can download forms from the site, including a link to Adobe’s site for the Acrobat Reader program.

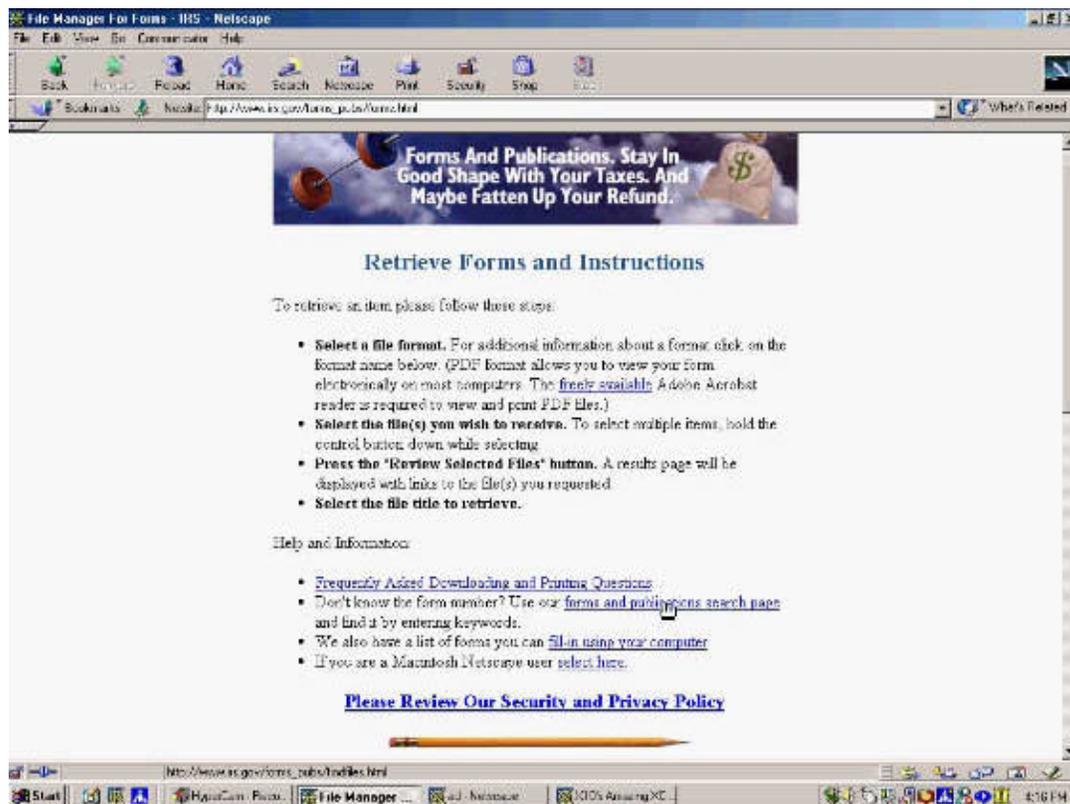
Moreover, there is also a link to a search page. On the bottom of this screen shot, one sees a pencil. As a 53-year-old librarian noted, “The pencil makes you think you’re at the end.” She was not the only respondent to have this impression. Several other users never made it below the pencil; rather, they clicked on the search button and got to the forms only after many more clicks. Dozens of users who eventually made it to this page hesitated in confusion and spent as much as 15-20 seconds reading the instructions, before scrolling down

FIGURE 5: SCREENSHOT OF AN IRS PAGE (FALL, 2001)



to notice that they were already on the right page to download the desired tax form.

Ironically, although the IRS updated its Web site in 2002 and removed the pencil design, the forms section is still not visible “above the fold” on the Forms and Instructions page (see Figure 7 for an example of what today’s page looks like). One obvious way to improve user access would be to place the forms on top of the page and then have a visible link to Instructions next to it. Putting instructions first—instructions that are in fact irrelevant to many of the users who already have Adobe Acrobat installed—confuses users because they cannot see that they have reached the target page, thus wasting their time as they read through irrelevant instructions.

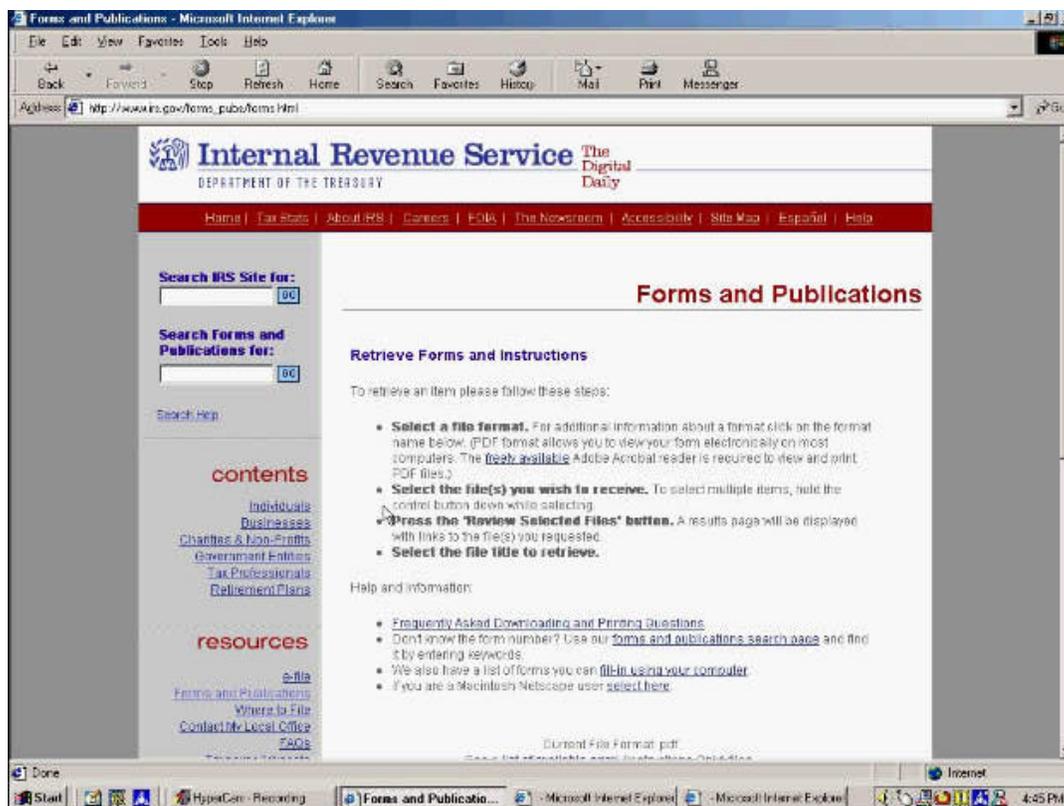
FIGURE 6: SCREENSHOT OF THE IRS'S FORMS PAGE (FALL, 2001)

CONCLUSION

Based on the search patterns of 97, randomly selected Internet users, one can see that users look for government forms in different ways. For something as widely known as the Internal Revenue Service, a significant number of people will try to access the agency's site directly. Although one may hypothesize that those who know the exact URL of an agency will more quickly access the desired information, if the site is not well-organized and it contains confusing design elements, then directly accessing the site will not make a user more efficient in obtaining the information sought. In fact, there is no statistically significant difference in time-to-completion between those who went to www.irs.gov as their first move in this search vs. those who tried elsewhere. This is probably due to some of the confusion that results from site organization, layout and design.

There are several obvious ways to improve the IRS website. First, site designers need to be more conscious of the consequences of their design decisions. In particular, search forms need to be available directly on the welcome page of a site. As shown in Figure 8, the 2002 version of the IRS's site

FIGURE 7: SCREENSHOT OF THE IRS'S FORMS AND INSTRUCTIONS PAGE (SPRING, 2002)



has a Forms and Publications search button, to which several of the respondents from the 2002 portion of the study turned immediately upon arrival at the site. Moreover, the focus of a page—in this case the link to the actual forms—should be made clearly visible on the section of the page that will be on the screens of users with modest-sized monitors (15” or 17”).

Second, a Web site should be designed so that when a user types in its address without “www,” the user will be redirected to the “www” version of the address. Many sites have implemented this usability feature. Why are there still sites that do not assist users in this way?

Third, users need to be educated about the importance of domain-name extensions, so that they are not derailed by commercial sites that do not necessarily have in their interest to point searchers quickly to the information they probably seek.

FIGURE 8: SCREENSHOT OF WWW.IRS.GOV (SPRING, 2002)



Following just these few recommendations would have eliminated the majority of problems these users encountered in searching for tax forms in this study. It would have shortened the amount of time they spent on the task, and it would probably have resulted in notably less frustration on their parts.

REFERENCES

- Benton, Foundation, and Leadership Conference on Civil Rights Education Fund. 2002. "Bringing a Nation Online: The Importance of Federal Leadership." Washington, DC: Joint Report.
- Bucy, E. 2000. Social Access to the Internet, *Harvard International Journal of Press/Politics* 5(1) p.50–61.
- Compaine, B. 2001. "Declare the War Won." In B. Compaine (ed.). *The Digital Divide: Facing a Crisis or Creating a Myth?* Cambridge, MA: MIT Press. p. 315-335.

- DiMaggio, P., Hargittai, E., Neuman, R. and Robinson, J. 2001. Social Implications of the Internet, *Annual Review of Sociology* 2, p. 307–336.
- Hargittai, E. In press. Beyond Logs and Surveys: In-Depth Measures of People's Web Use Skills, *Journal of the American Society for Information Science and Technology Perspectives*.
- Hoffman, D. L. and Novak, T. P. 1998. Bridging the Racial Divide on the Internet, *Science*. Vol 280. p. 390-391. Howard, P. E., Rainie, L. and Jones, S. 2001. Days and Nights on the Internet: The Impact of a Diffusing Technology, *American Behavioral Scientist*, 45(3) p. 383–404.
- Jansen, B.J. and 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(3) p. 235–246.
- McDonald, S. and Spencer, L. 2000. Gender Differences in Web Navigation: Strategies, Efficiency, and Confidence. In E. Balka and R. K. Smith (eds.). *Women, Work, and Computerization: Charting a Course to the Future*. Boston: Kluwer Academic Publishers.
- National Telecommunications and Information Administration. 2002. *A Nation Online*. Washington, DC.
- Panicware. 2001. "Don't Panic 4.0."
- Strover, S. 1999. *Rural Internet Connectivity*. Columbia, MO: Rural Policy Research Institute.
- Wang, P., Hawk, W.B. and Tenopir, C. 2000. Users' Interaction with World Wide Web Resources: An Exploratory Study Using a Holistic Approach, *Information Processing and Management* 36(2) p.229–251.
- Wilson, E.J. 2000. "Closing the Digital Divide: An Initial Review: Briefing with the President." Washington, DC: Internet Policy Institute.

¹ No default page was set on browsers in order not to influence respondents' initial actions once online. The sessions were started off by the researcher asking the respondent to recall, if possible, the default homepage on the computer she uses the most.

² The Hypercam program from Hyperionics was used on the PC and SnapZPro program from Ambrosia Software Inc. was used on the iMac. Captured items included the whole screen, every user action—e.g., mouse clicks and scrolling—and every comment the respondent made during the search.

³ The author conducted 77 of the interviews; two research assistants administered the remaining twenty.

⁴ All screen shots are from actual study sessions.