Succinct Survey Measures of Web-Use Skills

Eszter Hargittai and Yuli Patrick Hsieh
pubs@webuse.org and yulihsieh2012@u.northwestern.edu
Northwestern University

This is pre-print version of the paper forthcoming in Social Science Computer Review:

Hargittai, E. & Hsieh, Y.P. (2012). Succinct Survey Measures of Web-Use Skills. *Social Science Computer Review.*

http://webuse.org/p/a34

Abstract

There is a dearth of survey instruments for measuring Internet skills. This paper presents results from additional implementations of a previously-developed index measure. It considers the performance of the original instrument over time as well as shortened versions of it on two surveys of different populations. Drawing on analyses of five different data sets, the paper makes recommendations for various length survey items for measuring people's Web-use skills.

Keywords: Internet, Web use, skills, digital literacy, digital divide, digital inequality, surveys, methodology, measures

Acknowledgements

The authors appreciate the generous support of the John D. and Catherine T. MacArthur Foundation for making possible the main project upon which this paper draws. They are also grateful to Noshir Contractor for access to his data set and John Horrigan for his work on the FCC data set. The authors thank Ron Anderson and the anonymous reviewers for their very helpful comments. They appreciate the assistance of Waleeta Canon, Gina Walejko, Ericka Menchen–Trevino, Jessica Diamond, Cassi Saari and the rest of the undergraduate and graduate research assistants in the Web Use Project group from 2007-2010 with data collection and data entry. The first author also thanks the Center for Advanced Study in the Behavioral Sciences at Stanford and Harvard's Berkman Center for Internet & Society for enabling work on this project. The support of Ann Feldman, Tom Moss, and Karen Mossberger is also kindly acknowledged.

Introduction

As the Internet has spread to an increasingly large portion of the population, scholars have called for more nuanced measures of its uses including people's abilities to use the medium effectively and efficiently, i.e., their Web-use skills (e.g., Hargittai 2002, Hargittai 2010, DiMaggio et al 2004, Gui & Argentin In Press, Livingstone & Helsper 2010, van Dijk 2005, van Deursen and van Dijk 2009). A significant challenge in this domain has been the dearth of reliable instruments to measure people's online know-how. Some work has developed nuanced measures using in-person observations (e.g., Hargittai 2003, van Deursen and van Dijk 2009) offering detailed information about how people navigating the Web. However, due to the cost and labor associated with such methods, they are extremely difficult to replicate on more generalizable and larger samples leaving a need for survey instruments to capture information about people's Web-use skills.

In previous work, Hargittai (2009) suggested the use of a list of items to measure people's Internet skills based on the results and expansion of a study that compared people's actual online abilities with their responses to survey questions about Internet know-how (Hargittai 2005). The proposed list includes 27 Internet-related terms of which respondents are asked to rate their level of understanding on a 1-5-point scale. While useful, the recommendation poses a challenge to those working on surveys whose focus is not on Internet skills per se thus having less space for related questions. To address this potential concern with the original instrument, this research note draws on subsequent implementations and analyses of the 27-item index to offer suggestions for indexes that have fewer components yet are optimal in capturing variation among respondents' Web-use skills.

Brief review of the development of the index

As explained in the paper on "Survey Measures of Web-Oriented Digital Literacy" (Hargittai 2005), an initial index of seven items was the result of a study that included data both on people's

actual as well as their self-reported skills. In that study, 100 randomly-sampled people from a New Jersey county performed online information-seeking tasks at the study location resulting in measures of actual skill. That project, conducted in 2001 and 2002, yielded data about respondents' ability to perform online information-seeking tasks as well as how long they took to do so. The project had also asked respondents their level of understanding of over 40 Internet and computer-related terms. Of these, Hargittai (2005) identified an index of seven that correlated more highly with measures of actual skill than did more traditional proxies for skill such as time spent online, number of Internet use years and Internet self-efficacy (p.376.).

Due to the fast-changing nature of Internet tools and services, it is important to update such an instrument over time. In 2007, Hargittai (2009) collected data on respondents' understanding of an expanded list of items reflecting changes in the online landscape. In this study, 1,189 diverse young adults rated their understanding of 30 Internet-related terms. These items were presented on the survey in two groups, fifteen terms each. Three of the 30 items were bogus terms (i.e., JFW, proxypod and filtibly) to test whether respondents were randomly checking off responses. As noted in the paper (p.133-134), the bogus terms performed lowest on their respective lists suggesting that people's reports on these items were not random. That is, while there were no corresponding measures of actual skills in this study to test whether people's self-reported ratings correlated with their true abilities, nonexistent terms should yield the lowest rankings and indeed they did. The 27-item scale and shorter versions of it have since been replicated in other studies. Here, we draw on the results of four subsequently administered surveys – two conducted by Hargittai, two by others – to consider what may be possible shorter item lists for measuring online skills.

Recent implementation of the full instrument

With constant changes in online technologies, tools and services, it may be that certain measures become outdated while others rise in prominence and importance to people's daily online

activities. To test this proposition, we draw on data collection since 2007 – the year the original data set with 27 items had been collected – to illustrate the robustness of the overall measure over time. Hargittai (2009) drew on data collected in 2007 from the first-year cohort at the University of Illinois, Chicago (UIC)[1], a socioeconomically and racially diverse urban research university (US News & World report 2006). Over one thousand (1,060) first-year students took the paper-pencil survey administered in class for an 81.9% participation rate of all students enrolled in the course. The instrument included an item to verify students' attentiveness to the survey: "The purpose of this question is to assess your attentiveness to question wording. For this question please mark the 'Very often' response." A small portion of students, 3.4%, responded incorrectly suggesting that they were randomly checking off responses and thus were excluded from the analyses. In sum, 1,060 first-years answered the verification question correctly. Of these respondents, 1,004 provided useable responses for all of the knowledge items. They are the basis for the numbers we present below. We refer to this data set as the UIC 2007 survey. The first column under the heading UIC 2007 in Table 1 presents the mean of individual items in decreasing order from the survey reported in Hargittai (2009), the second column lists the item-rest correlation, which is the correlation of the item to the rest of the scale without that item.[2]

Hargittai replicated that same 27-item instrument on a similar survey in 2009 on the then entering cohort at UIC, i.e., a different group of students from those surveyed two years earlier. This survey had an overall response rate of 80.5% and also included the verification question mentioned above. The 4.5% of respondents who answered it incorrectly were excluded from the data set

[1] The Principle Investigator of this study (Hargittai) is not now nor has ever been affiliated with the University of Illinois, Chicago, other than in the context of this research project. The campus was chosen as the research site for the overall project due to the diversity of its student body and the existence of a class in the curriculum that all students are required to take.

[2] These figures are not the same as the ones reported in Hargittai (2009), because the numbers in that paper are based on 1189 respondents rather than focusing on just first-year students and those who had no missing values on any of the skill items.

resulting in 1,115 participants. Of these, 1041 had no missing values on the skill items and they constitute the basis for the numbers presented here. The columns in Table 1 under the heading UIC 2009 report the mode, mean, standard deviation and item-rest correlation of the items from this second study.

In 2010, Hargittai administered a follow-up survey on students who participated in the 2009 study gathering information about their ratings of the 27 items' understandings in that subsequent year. This study had a 45.3% response rate with 505 students from the original 1,115 participating. Like the previous UIC surveys, this one also included a verification question to assess attentiveness to question wording; 3% of the returned instruments did not give a correct answer and were thus excluded from the analyses. The 505 respondents had all filled out the verification question correctly and 483 of them had no missing values on the 27 skill items. The final group of columns in Table 1 presents the mode, mean, standard deviation and item-rest correlation for each item in this data set.

As the figures demonstrate, there is considerable consistency over time in the relative ranking of the items on the scale. Comparing the mean scores of the 2007 sample with those of the 2009 and 2010 samples, we see that the order of understanding on all surveys is highly similar. Additionally, there is great resemblance between the descriptive statistics of the same items in the 2009 sample and those in the 2010 sample, suggesting that our instrument performs consistently over time. This is also supported by the fact that the correlation of the mean skill score for 2009 and 2010 is .75 ($p<0.000$) among those who took the survey both years.

The consistency in the rankings of the terms also indicates that our full list of Web-related items reflects various levels of know-how among respondents. While many terms may be relatively easy and respondents reported high scores for them across surveys in different years, we also find that participants in the most recent survey continue to report a limited understanding of some of the lowest-scored items from the first survey. For example, as shown by the first rows in Table 1,

browser-related terms such as "reload", "favorites" and "bookmark" in addition to "advanced search" ranked the highest consistently across time and across samples. Given that using browsers is a common way of going online, it is not surprising that respondents have a high-level of understanding of related terms. However, respondents in different years consistently reported a limited understanding of terms like "bookmarklet", "cache", "widget", "phishing", "malware", "social bookmarking" and "RSS" suggesting that these concepts have remained elusive to many Web users over time.

The comparison across different years also suggests, however, that some terms move up on the ranking list as time passes. Some of the initially lower-scored items (i.e., "tagging", "tabbed browsing" and "wiki") reported in Hargittai (2009) based on the UIC 2007 survey became higher-scored items in the 2009 and 2010 surveys, suggesting that these Internet-related terms have become better-known than they had been in earlier years. These changes likely reflect trends in the increasing popularity of certain specific Internet services such as photo tagging on social network sites like Facebook, tabbed browsing functionality introduced to many Web browsers, and the popularity garnered by the site Wikipedia. In contrast, other terms such as "cache", "phishing", "social bookmarking" and "RSS" have occupied a stable position on the list as barely-understood terms.

Worthy of note is that the reliability tests of the 2009 and 2010 surveys also indicate a stable and consistent performance of the global Web-use skill index measure. The Cronbach's alpha values are .94 for the 2009 survey responses and .93 for the 2010 survey data both of which compare favorably to .94 of the UIC 2007 data set reported in Hargittai (2009). All of these figures are very high and suggest that the instrument is internally consistent. We also looked at item-rest correlations to determine the correlation of the item with the rest of the scale were that item excluded. The final column for each survey in Table 1 presents these figures. They suggest that there are no particularly weak items in the scale. Additionally, we also looked at what the alpha

values would be were the individual items removed from the scale. All alphas in all three data sets remain very close to the full scale measure, i.e., with less than a 0.01 change in all cases (due to space constraints and given the slight change in figures, these numbers are not included in the table).

While the full index is helpful, it is important to recognize that few surveys have room for such a large number of items. To address this concern, we now turn to a discussion of shorter alternatives for surveys with less room for such questions. In the next section, we highlight some important considerations for ensuring high-level performance while reducing the length of the instrument followed by two examples of shorter index implementations and our overall recommendations for shorter item lists.

Considerations for the development of shorter item lists

As noted in the section above, the 27 Internet-related terms reflect different understanding levels. While most items show a consistent level over time, we do find that people's reported level of understanding of some items may change due to increasing popularity of certain online services. The implication of this finding is that the range of levels at which people report understanding the items is a crucial factor when adopting this instrument of Web-use skills to conduct research on people's digital know-how. To ensure that the instrument captures variations in skill in any particular population, researchers have to include items that are understood at different levels even when using a shorter list.

Based on the mean, mode and median[3] of the original 27 terms, we identified three general levels of understanding: high, medium and low. We consider a term as a high-level understanding item if its mode and median are consistently of the higher scores (i.e., 4 or 5) and the majority of its mean scores from the different UIC survey data are higher than 4. In contrast, we consider a term as

[3] Due to space constraints, the tables in this piece do not include medians, but were consulted in the preparation of the suggested indexes. For those interested in the medians of the individual items in the UIC studies, they are available from the authors. [TO EDITOR: We could also put these up on our Web site and indicate that in the paper.]

a low-level understanding item if its mode and median are consistently of the lower scores (i.e., 1 or 2) and the majority of its mean scores from different UIC survey data are lower than 2.5. The terms that do not fall into either of these two categories are of medium-level understanding.

Another important criterion for selecting items concerns the characteristics of the population under study. If researchers were to survey people with less online experiences, then it makes more sense to include more of the highly-understood items than the ones that yield low levels of understanding given that such people are even less likely to understand terms that more experienced Internet users already find harder to comprehend. Likewise, if the targeted population is likely to have better online abilities (such as technical professionals), then researchers may want to include more of the less understood items instead of ones highly understood by most in order to capture variation within the group under investigation and to measure respondents' digital skills accurately. Beyond ensuring sufficient variation in the skills measure, adjusting the instrument based on the characteristics of the targeted population may also help researchers reduce non-response due to the heavy burden of survey questions (i.e., less knowledgeable respondents faced with numerous lesser-known items may abandon the instrument altogether in frustration).

Based on these considerations, we have worked with other projects to make recommendations for shorter item lists. In the next section, we present two cases of shorter lists having been implemented on samples different from our own. One study draws on a nationally-representative sample of Internet users; the other implements the survey on a low-income urban population.

The Web-use skill measure index applied in other studies

To demonstrate the utility of the Web-use skill index with fewer items, we present descriptive statistics of related constructs used in two other studies.

Study 1: The Federal Communications Commission's National Consumer Broadband Service Capability Survey

Study background

As part of the National Broadband Plan, the Federal Communications Commission conducted a national telephone survey in Fall 2009 to understand the status of broadband adoption and how Americans use the Web in their everyday lives (Horrigan 2010). The survey covered a broad range of questions, including types and places of Internet access, service costs, reasons of broadband adoption/non-adoption, online activities, attitudes about the Internet, and Web-use skills based on the instrument published in Hargittai (2009).

Data and methods

The National Consumer Broadband Service Capability Survey (henceforth the FCC Broadband Survey) was administered in English and Spanish over the phone with a nationally representative sample of 5,005 American adults in October and November 2009 (Horrigan 2010). The sampling frames consisted of a random-digit dial (RDD) landline and RDD cell sample (Horrigan 2010: 46). The researchers over-sampled non-adopters and employed a series of sample adjustments and weighting to ensure the national representativeness of the survey sample. For the purposes of this research note, we look at those Internet users in the sample who had no missing values on the skill measures constituting 3,121 participants.

As part of the section looking at people's online activities and Internet uses, the survey included questions asking respondents about their knowledge of six Web-related terms. It is important to note that the wording and format are somewhat different from Hargittai's (2009) skills measure, nonetheless, the study offers a helpful case of implementing a shorter skill item list on an instrument. The response category of the FCC Broadband survey is a 4-point scale rather than a 5-point scale. The Web-related terms in the survey are: (1) internet browser cookie; (2) spyware and

malware; (3) operating system; (4) refresh or reload; (5) widget; and (6) JPEG file. The reliability test indicates that these six items in the FCC data set are consistent with an alpha value of .88.[4]

Performance of Web-use skill items

There are differences in question wording and the response scale between the FCC Broadband Survey and the instrument used in the UIC studies so the descriptive statistics of the FCC Broadband Survey's six items are not directly comparable to the corresponding terms used on the UIC surveys. They are nonetheless instructive. The figures presented in Table 2 show basic descriptive statistics for the six items on the FCC Broadband Survey for those respondents who are Internet users in the sample. It is clear that "widget" is the least recognized term from among the six included items. The statistics of the other 5 items are located at the higher-end of the scale with "reload or refresh" as the most understood term.

It is important to point out that one of the six questions about know-how on the FCC Broadband Survey contained two different terms from the original 27-item list; "spyware and malware". In the UIC surveys, these two items behaved rather differently when asked separately thus grouping them into one item is not ideal. While participants in the UIC surveys rated their level of understanding of "spyware" relatively highly, they exhibited a considerably lower level of understanding of "malware". Accordingly, results of that particular question from the FCC Broadband Survey must be interpreted with caution. Future surveys that include these terms should do so as separate items.

Including more items from the middle range of the original list may have led to more nuanced variation on the final skill index of the FCC Broadband Survey sample. Nonetheless, even

[4] The FCC Broadband Survey data are available to the public at http://www.fcc.gov/broadband-consumer-survey/Public-Posting.zip. The figures we report in this piece are ones we calculated on the data set ourselves.

with these six items representing relatively well understood terms, the index provides helpful variation in the Web-use skills of respondents.

Study 2: North Kenwood/Oakland area survey of the Chicago Climate Action Plan

Study background

The Chicago Climate Action Plan (CCAP) is an initiative developed and sponsored by a multi-stakeholder task force of the City of Chicago that aims to describe the implications of climate change on Chicago and to encourage practices that address the challenges posed by it. With the goal of engaging diverse communities in the CCAP, Northwestern University's Science of Networks in Communities (SONIC) Research Lab partnered with Chicago's Field Museum to conduct a social network survey in several Chicago communities in order to understand to whom and to which organizations as well as to what media the residents from these areas would turn for information about environmental issues. Another main objective of the study is to provide the task force with recommendations to promote greater awareness of and engagement in the CCAP across diverse communities through the use of social networks. The first community surveyed as part of this project was the North Kenwood/Oakland (NKO) area in the south side of Chicago.

Data and methods

Through computer-assisted personal interviewing, the SONIC research team collected data in the NKO area between November 2009 and February 2010. Given the exploratory nature of the project, the researchers employed both quota and convenience sampling strategies to recruit participants. The demographic quota was constructed based on the 2000 Census data of the NKO area and interviewers were instructed to use such information for recruitment. Participants were recruited by interviewers at public places in the community such as cafés, supermarkets and community centers; the survey could also be administered in private residences if necessary.

On certain demographic characteristics (i.e., gender and race) the survey sample is representative of the area's make-up while on others there is some level of over or underrepresentation (i.e., people with a college degree and those with household incomes higher than $60,000 are overrepresented, while people with some college experience are underrepresented).[5] The survey resulted in a total of 218 valid responses. As part of the question module looking at information networks of peer residents, the research team adopted Hargittai's Web-use skill measure in order to understand whether people's online abilities may be related to their media uses for communicating with their peers about environmental information. Due to questionnaire length limitations, the research team only included 15 Internet-related items in the NKO survey instead of the original 27 proposed by Hargittai (2009). Based on the distribution of reported know-how in the UIC sample, the researchers included nine highly-understood Web-related terms and six items that were understood at much lower levels, some with large variations, others with smaller variance. The CCAP-NKO survey included the same question wording and response categories as Hargittai's study. Having worked with Hargittai to run some analyses on the original UIC 2007 data set, the CCAP-NKO researchers were able to establish that the 15 items were still in considerable agreement with an alpha value of .90 based on responses to the UIC 2007 survey.

Performance of Web-use skills items

The descriptive statistics in Table 3 represent the responses of the 161 Internet users in the sample who had no missing items on the skill measures. The figures indicate that the responses of the 15 Internet-related items are consistent with the difficulty levels identified in the UIC 2007 sample reported in Hargittai (2009). Like UIC respondents, NKO survey participants also reported having better knowledge of the terms with higher mean scores in the original surveys. The reliability

[5] SONIC researchers were able to establish these comparisons by extracting the population data of the Census tracts that consist of the NKO area from the 2000 U.S. Census through the U.S. Census Bureau's Web portal – American FactFinder (http://factfinder.census.gov).

test of these 15 items have a high alpha value (alpha=.97) in the NKO sample, suggesting that the shortened skill items remain highly consistent. The standard deviations of individual items are also very similar to the statistics of the corresponding items on the UIC surveys.

Perhaps not surprisingly, given that many NKO residents come from lower socioeconomic backgrounds, which is related to lower levels of Internet use (e.g., Mossberger, Tolbert and Stansbury 2003; Stern, Adams and Elsasser 2003), which in turn may influence Internet know-how, NKO participants reported having a lower average composite score than UIC participants. In addition, the results show that two of the items understood well by participants in other studies (i.e., "advanced search" and "firewall") display a bi-modal distribution with the lowest score being one of the most frequent responses. Additionally, in this sample, 8 of the 15 items have a mean score lower than 3 (i.e. the average point of the 5-point scale). The implication of these findings is that the Web-use skill measure performs well beyond a population of college students for capturing variation in online know-how.

Recommendations for shorter item lists

Based on the case studies of shorter item lists' empirical implementations presented above as well as the analyses of different surveys administered at UIC, here we make recommendations for shorter lists of items that still perform well when included on a survey. We provide two lists; one to be included on surveys administered to the *general population* (Table 4a), the other to people *with lower levels of online experiences* (Table 4b). First, on Tables 4a and 4b, we present a list consisting of six items followed by a list of 10 and then 15 items. The higher the number of items, the more reliable the scale, but even shorter scales demonstrate high consistency as noted in the bottom section of the table with Cronbach's alpha figures.

When constructing the list, it is also important to consider the relevance and applicability of the terms to different contexts. For some of the terms in the original 27-item list, users may have to

engage in specific online practices, such as using information aggregators (i.e., "RSS" or "Web feeds") or file sharing (i.e., "torrent") in order to become more knowledgeable about them. While those online activities may be popular with some groups, they are not necessarily of interest to all Web users or relevant to every user's needs. Therefore, a shorter instrument with too many such specific terms may bias against users who do not engage in very concrete types of online activities. To this end, especially for the shorter lists, our emphasis is on selecting terms related to fairly general uses (e.g., browsers, common file types) as well as Internet security and privacy matters given that all Web users are subject to risks of identity theft and other online security threats regardless of their preferences among particular online activities.

On the six-item list, we include two high-level understanding items (i.e., "advanced search" and "PDF"), two medium-level items (i.e., "spyware" and "wiki") and two low-level items (i.e., "cache" and "phishing") for surveying the general population. For an instrument targeting less-experienced populations, we replace "cache" with "preference settings" in order to reduce the overall difficulty level of that list. While only one-third of the items on the list for a survey of the general population consists of highly-understood items, half of the items on the list for disadvantaged populations are such terms.

With the longer list of items, our goal is to maintain the proportion of each difficulty level to the extent possible. The proportion of high-level understanding items in the 10-item and 15-item lists for studying general populations is 33% and 40% respectively. Likewise, 40% of the terms in the 10-item and 15-item lists for surveying less-experienced populations are high-understanding-level items. While we selected "tagging" and "tabbed browsing" as a part of the high-level understanding items for the longer list targeting the general population, we included the browser-related terms "favorites" and "bookmark" for less-experienced populations. Were researchers to change individual

items, our recommendation is to select others from the list at the same level of understanding as the ones they are replacing.

Despite rapid changes in technology, the understanding of terms exhibits considerable consistency across the years. That said, as new tools and services emerge and are incorporated into mainstream Internet uses, the testing and addition of new items will be necessary. When considering additions, researchers should avoid choosing terms that are platform-specific (e.g., terms that are more common with the use of Windows or Mac operating systems) so as not to bias toward users of particular services unless examining the know-how of users of a specific system is the goal of the investigation.

Reliability tests from prior UIC surveys suggest that all of our recommended lists are in good agreement as per the Cronbach's alpha values listed on the bottom of the table for each proposed index. Not surprisingly, the longer the list, the higher the alpha value and thus when possible, the inclusion of more items is better.

Conclusion

Our goal has been to analyze the performance of a previously developed Web-use skill index survey measure and make recommendations on abbreviated versions for instruments of various lengths and for different targeted populations. Recent empirical implementations of abbreviated instruments suggest that shorter online skill indexes still work reliably and consistently as compared to the longer version. We also demonstrate the utility of the shorter instruments by showing that the selected items in every recommended list are in good agreement with the original 27-item instrument from all three UIC surveys. The use of more items results in higher alpha values, however, and thus we recommend the inclusion of as many components as possible.

Table 1: Descriptive statistics of Web-use skill measures in the UIC 2007, UIC 2009 and UIC 2010 surveys

	UIC 2007 survey		UIC 2009 survey				UIC 2010 survey			
Item	Mean	IRC*	Mode	Mean	SD	IRC	Mode	Mean	SD	IRC
Reload	4.69	0.45	5	4.75	0.59	0.47	5	4.74	0.60	0.43
Advanced search	4.47	0.46	5	4.41	0.91	0.53	5	4.39	0.93	0.54
Favorites	4.43	0.49	5	4.40	0.94	0.55	5	4.46	0.87	0.52
Bookmark	4.40	0.49	5	4.42	1.00	0.50	5	4.45	0.90	0.51
Spyware	4.11	0.58	5	3.83	1.23	0.63	5	3.48	1.28	0.67
Preference setting	4.03	0.61	5	4.03	1.05	0.59	5	3.93	1.10	0.64
Blog	3.98	0.54	5	3.72	1.30	0.55	5	3.57	1.32	0.62
Firewall	3.84	0.66	5	3.76	1.22	0.73	5	3.57	1.23	0.70
PDF	3.57	0.60	5	4.01	1.14	0.58	5	4.08	1.00	0.57
JPG	3.42	0.65	5	3.32	1.47	0.63	4	3.45	1.34	0.58
Tagging	3.35	0.44	5	4.09	1.26	0.42	5	4.42	1.05	0.26
Weblog	3.27	0.60	3	3.20	1.36	0.60	3	3.05	1.38	0.61
Newsgroup	2.92	0.65	3	2.74	1.25	0.60	2	2.53	1.29	0.60
Tabbed browsing	2.79	0.64	5	3.78	1.51	0.52	5	4.00	1.43	0.45
Frames	2.76	0.72	1	2.48	1.34	0.67	1	2.44	1.36	0.64
Podcasting	2.74	0.66	1	2.94	1.51	0.67	1	2.74	1.44	0.64
Web feeds	2.54	0.69	1	2.76	1.50	0.71	1	2.83	1.42	0.64
Torrent	2.44	0.67	1	2.85	1.71	0.64	1	3.01	1.65	0.57
Bcc (on-email)	2.42	0.53	1	2.50	1.52	0.53	1	2.79	1.59	0.46
Bookmarklet	2.33	0.52	1	2.27	1.29	0.52	1	2.14	1.30	0.50
Wiki	2.28	0.62	5	3.46	1.57	0.60	5	3.74	1.35	0.57
Cache	2.28	0.70	1	2.41	1.56	0.72	1	2.48	1.50	0.66
Widget	1.79	0.56	1	2.36	1.53	0.60	1	2.45	1.52	0.63
Phishing	1.78	0.60	1	2.12	1.45	0.62	1	2.07	1.39	0.62
Malware	1.75	0.64	1	2.20	1.54	0.63	1	2.62	1.51	0.64
Social bookmarking	1.68	0.61	1	2.05	1.34	0.63	1	2.26	1.38	0.59
RSS	1.61	0.65	1	1.84	1.26	0.65	1	1.83	1.24	0.62
Web-use skills (composite score)	3.03	-	-	3.21	0.84	-	-	3.24	9.79	-
N	1004		1041				483			
Scale	5-point		5-point				5-point			
Cronbach's alpha	0.9404		0.9413				0.9350			

* IRC refers to Item-Rest Correlation, see text for explanation. We only include respondents with valid responses for all 27 items in the analyses of both the UIC 2009 and UIC 2010 surveys.

** In Hargittai (2009), the sample size was 1189 as it included all respondents, not just first-year students. The analyses are based on the 1,004 first-year respondents in the sample who had no missing values on the 27 items.

Table 2: Descriptive statistics of Web-use skill measures on the FCC Broadband Survey

	FCC-Broadband survey* (Web-user sample)				UIC 2007
Item	Mode**	Mean	SD	IRC**	Mean
Refresh or reload	4	3.17	1.09	0.67	4.69
Operational system	4	2.90	1.11	0.73	-
Spyware and malware	4	2.87	1.10	0.72	4.11 and 1.75
Cookie	4	2.84	1.13	0.74	-
JPEG	4	2.59	1.28	0.71	3.42
Widget	1	1.78	1.10	0.52	1.79
Web-use skills (composite score)	-	2.69	0.89		-
N		3121			1004***
Scale		4-point			5-point
Cronbach's alpha****		.87			0.70

* We have reverse coded the original 4-point scale in order to make the scoring pattern show a consistent positive direction. Here, 4 refers to *very well*; 3 refers to *well*; 2 and 1 refers to *not too well* and *not at all* respectively. We only include respondents with valid responses on all items in the analyses.
** IRC refers to Item-Rest Correlations.
*** In Hargittai (2009), the sample size was 1189 as it included all respondents, not just first-year students. The analyses are based on the 1,004 first-year respondents in the sample who had no missing values on the 27 items.
**** The Cronbach's alpha value for the UIC 2007 survey measures is based on the reliability test of the following five separate items: "reload", "spyware", "malware", "JPG" and "widget." As noted in the text, "spyware" and "malware" were asked separately on the UIC surveys. Also, "operaiontal system" and "cookie" were not on the UIC survey.

Table 3: Descriptive statistics of Web-use skill measures on the CCAP-NKO survey

	CCAP-North Kenwood/Oakland area survey*				UIC 2007
Item	Mode	Mean	SD	IRC**	Mean
Favorites	5	3.366	1.448	0.80	4.43
Bookmark	4	3.267	1.482	0.80	4.40
Advanced search	1 and 4	3.261	1.473	0.82	4.47
Firewall	1 and 4	3.068	1.424	0.86	3.84
JPG	4	3.056	1.574	0.83	3.42
PDF	1	3.025	1.585	0.86	3.57
Preference setting	1	3.012	1.529	0.88	4.03
Spyware	1	2.963	1.466	0.88	4.11
Weblog	1	2.652	1.437	0.87	3.27
Newsgroup	1	2.652	1.411	0.82	2.92
Wiki	1	2.596	1.547	0.83	2.28
Podcasting	1	2.447	1.508	0.87	2.74
Phishing	1	2.385	1.5	0.84	1.78
Malware	1	2.261	1.506	0.81	1.75
RSS	1	2.037	1.387	0.73	1.61
Web-use skills (composite score)	-	2.80	1.27		-
N		161			1004***
Scale		5-point			5-point
Cronbach's alpha		.97			.90

* We only include respondents with valid responses on all items in the analyses.
** IRC refers to Item-Rest Correlations.
*** In Hargittai (2009), the sample size was 1189 as it included all respondents, not just first-year students. The analyses are based on the 1,004 first-year respondents in the sample who had no missing values on the 27 items.

Table 4a. Recommended lists of items for different size indexes to measure Web-use skills, for general populations

Full list of 27 items in decreasing order of understanding as per UIC 2009 results	Abbreviated Web-use skills indexes for the *general* population		
	6 items	10 items	15 items
Reload (H)*			
Bookmark (H)			
Advanced search (H)	Advanced search	Advanced search	Advanced search
Favorites (H)			
Tagging (H)		Tagging	Tagging
Preference setting (H)			Preference setting
PDF (H)	PDF	PDF	PDF
Spyware (M)*	Spyware	Spyware	Spyware
Tabbed browsing (M)			Tabbed browsing
Firewall (M)			Firewall
Blog (M)			
Wiki (M)	Wiki	Wiki	Wiki
JPG (M)		JPG	JPG
Weblog (M)		Weblog	Weblog
Podcasting (M)			Podcasting
Torrent (M)			
Web feeds (M)			
Newsgroup (M)			
Bcc (on-email) (M)			
Frames (L)*			
Cache (L)	Cache	Cache	Cache
Widget (L)			
Bookmarklet (L)			
Malware (L)		Malware	Malware
Phishing (L)	Phishing	Phishing	Phishing
Social bookmarking (L)			
RSS (L)			RSS
Reliability tests from prior UIC surveys (α)			
2007	0.77	0.85	0.90
2009	0.79	0.86	0.90
2010	0.78	0.84	0.90

* H=high-level understanding; M=medium-level understanding; L=low-level understanding.

Table 4b. Recommended lists of items for different size indexes to measure Web-use skills of people with low levels of Internet experiences

Full list of 27 items in decreasing order of understanding as per UIC 2009 results	Abbreviated Web-use skills indexes for populations with *low levels of Internet experiences*		
	6 items	10 items	15 items
Reload (H)*			
Bookmark (H)			Bookmark
Advanced search (H)	Advanced search	Advanced search	Advanced search
Favorites (H)		Favorites	Favorites
Tagging (H)			Tagging
Preference setting (H)	Preference setting	Preference setting	Preference setting
PDF (H)	PDF	PDF	PDF
Spyware (M)*	Spyware	Spyware	Spyware
Tabbed browsing (M)			
Firewall (M)			Firewall
Blog (M)			
Wiki (M)	Wiki	Wiki	Wiki
JPG (M)		JPG	JPG
Weblog (M)		Weblog	Weblog
Podcasting (M)			Podcasting
Torrent (M)			
Web feeds (M)			
Newsgroup (M)			
Bcc (on-email) (M)			
Frames (L)*			
Cache (L)			
Widget (L)			
Bookmarklet (L)			
Malware (L)		Malware	Malware
Phishing (L)	Phishing	Phishing	Phishing
Social bookmarking (L)			
RSS (L)			RSS
Reliability tests from prior UIC surveys (α)			
2007	0.76	0.85	0.89
2009	0.78	0.86	0.90
2010	0.79	0.86	0.89

* H=high-level understanding; M=medium-level understanding; L=low-level understanding

References

DiMaggio, P., E. Hargittai, C. Celeste, and S. Shafer. 2004. "Digital Inequality: From Unequal Access to Differentiated Use." Pp. 355-400 in *Social Inequality*, edited by Kathryn Neckerman. New York: Russell Sage Foundation.

Gui, M and G. Argentin. In Press. "Digital Skills of Internet Natives: Different Forms of Internet Literacy in a Random Sample of Northern Italian High School Students." *New Media & Society*.

Hargittai, E. 2002. "Second Level Digital Divide: Differences in People's Online Skills." in *First Monday*.

Hargittai, E. 2003. "How Wide a Web? Inequalities in Accessing Information Online." in *Sociology Department*. Princeton, NJ: Princeton University.

Hargittai, E. 2005. "Survey Measures of Web-oriented Digital Literacy." *Social Science Computer Review 23 (3)*:371-379.

Hargittai, E. 2008. "The Digital Reproduction of Inequality." in *Social Stratification*, edited by David Grusky. Boulder, Colorado: Westview Press.

Hargittai, E. 2009. "An Update on Survey Measures of Web-Oriented Digital Literacy." *Social Science Computer Review. 27(1):* 130-137.

Hargittai, Eszter. 2010. "Digital Na(t)ives? Variation in Internet Skills and Uses among Members of the "Net Generation"." *Sociological Inquiry* 80:92-113.

Horrigan, J. B. 2010. "Broadband Adoption and Use in America." in *The FCC Omnibus Broadband Initiative (OBI) Working Paper Series*. Washington DC: Federal Communications Commission.

Livingstone, L. and E. Helsper. 2010. Balancing opportunities and risks in teenagers' use of the internet: the role of online skills and internet self-efficacy. *New Media & Society*. 12(2):309-329

Mossberger, K., C. J. Tolbert, and M. Stansbury (2003). Virtual Inequality: Beyond the Digital Divide. Washington, DC, Georgetown University Press.

Stern, M.J., Adams, A.E. & Elsasser, S. 2009. "Digital Inequality and Place: The Effects of Technological Diffusion on Internet Proficiency and Usage across Rural, Suburban, and Urban Counties." *Sociological Inquiry* 79(4): 391-417

US News and World Report. 2006. "Campus Ethnic Diversity: National Universities." in *America's Best Colleges 2007*. Washington, DC.: U.S. News Media Group.

van Deursen, A. J. A. M., and J. A. G. M. van Dijk. 2009. "Improving digital skills for the use of online public information and services." *Government Information Quarterly* 26:333-340.

van Dijk, Jan A.G.M. 2005. *The Deepening Divide: Inequality in the Information Society*. London: Sage Publications.