2009

Does Web Site Usability Correlate with Web Site Usage?

Charles E. Downing
Northern Illinois University

Chang Liu
Northern Illinois University

Follow this and additional works at: https://scholarworks.lib.csusb.edu/jitim

Part of the Management Information Systems Commons

Recommended Citation
Available at: https://scholarworks.lib.csusb.edu/jitim/vol18/iss3/10

This Article is brought to you for free and open access by CSUSB ScholarWorks. It has been accepted for inclusion in Journal of International Technology and Information Management by an authorized editor of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.
Does Web Site Usability Correlate with Web Site Usage?

Charles E. Downing  
Northern Illinois University  
USA

Chang Liu  
Northern Illinois University  
USA

ABSTRACT

This study compares usability ratings of Web sites, based on academic research, to actual usage data, from Nielsen Online. This is accomplished by consulting academic usability research, and using those guidelines to rate eight retail Web sites on usability. Then, this study uses “real world” data, the usage data from Nielsen Online, to longitudinally validate the academic findings. Results show that, over time, the calculated usability ratings move in the same direction and in the same magnitude as the usage data.

INTRODUCTION

Electronic commerce (EC) has allowed business organizations to enhance their economic growth, improve efficiency and effectiveness, keep inventories lean, and reduce cost. More and more companies go online to conduct business transactions using the Internet and Web technology. Research indicates that EC will continue to grow, and the number of Internet users worldwide reached over 1.57 billion as February 2009 (http://internetworldstats.com, 2009). Thus, more than one fifth of the total world population uses the Internet. And this penetration is rapidly increasing, with broadband connections having grown from 55% in early 2008 and 47% in 2007. Indeed, EC is now a reality of our daily lives. Intuitively, companies need to ensure that their Web presence is accessible and acceptable to this huge market to maintain future profitability.

Usability is one prominent measure of accessibility and acceptability. Usability has been studied extensively in both the human-computer interaction discipline (e.g., Palmer, 2002; Këpuska, Gurbuz, Rodriguez, Fiore, Carstens, Converse, & Metcalf, 2008) and in the electronic commerce area (Agarwal & Venkatesh, 2002; Eighmey & McCord, 1998; Massy, Khatri, & Montoya-Weiss, 2007; Nielsen, 2000; Palmer & Griffith, 1998; Venkatesh & Agarwal, 2006; Venkatesh, Morris, Davis, & Davis, 2003). Although the Web provides an opportunity for a firm to offer a unique and satisfying experience to its visitors, developing a user friendly Web site is not a simple task. One of the critical challenges facing businesses today is to develop a Web site that is not only compelling for the visitors, but is also able to serve the goals of the business as well. A major business goal of EC is Web site usage, which can be measured by number of visitors, time per visitor, and page views a site receives. None of the research on Web site usability has examined the relationship between what is deemed a usable site and usage of that site. The purpose of this research is to rate Fortune 500 retail Web sites’ usability given previous guidelines, and then cross-reference whether the usability of the sites, as rated by survey
respondents, correlates with usage of the sites. The research will present empirical usability results obtained from users of operating retail Web sites, and will compare that to Web site usage data from Nielsen Online.

BACKGROUND

Usability of a Web site is a characteristic business professional and academic researchers have long realized to be important (Agarwal & Venkatesh, 2002; Massy et al., 2007; Palmer, 2002; Post, Kagan, & Sigman, 2009; Venkatesh & Agarwal, 2006; Venkatesh & Ramesh, 2006). Usability has been conceptually defined and operationally measured in different ways. This paper adopts the ISO 9241 definition of usability – “The extent to which a product or a service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” The purpose of usability research is to understand how to design and engineer more usable Web sites in an attempt to identify a set of principles and common practices that will ensure usability is an outcome of system design. In the context of Web site design, Nielsen (2000) claimed that usability engineering is the key to successfully conducting commercial Web site design and argued that a business firm that does not pay enough attention to this aspect is not likely to generate sustainable traffic to its Web site or, even if visitors stop by, the site will lack the crucial attribute of “stickiness”, or usage, which often translates into customers.

Therefore, as mentioned, the topic of usability has been studied in both human computer interaction literature (e.g., Palmer, 2002) and information systems literature (e.g., Green & Pearson, 2009; Venkatesh & Agarwal, 2006). Prior research has proposed multiple factors by which the usability of Web sites may be assessed. Eightmey and McCord (1998) derived seventeen factors tapping into different aspects of Web site usability that were subsequently reduced to nine groups including dimensions such as personal involvement, useful information, simplicity of organization, and desire for relationship. Agarwal and Venkatesh (2002) employed the Microsoft Usability Guidelines (MUG) to assess usability of multiple Web sites from four different industry sectors: airlines, online bookstores, automobile manufacturers, and car rental agencies. The MUG guidelines include five major categories that collectively tap into different aspects of Web site usability: content, ease of use, promotion, made-for-the medium, and emotion. Venkatesh and Ramesh (2006) concluded that the MUG-based model outperformed the widely employed Technology Acceptance Model (TAM) (see, e.g., Holsapple and Wu, 2008; Mahatanankoon, Klaus, & Wen, 2007) both in terms of richness and variance explained (about 70 percent compared to 50 percent). Moreover, Web site usability is a significant antecedent of purchase behavior (Venkatesh and Agarwal, 2006). An excellent summary of usability studies can be found in the research of Green & Pearson (2009).

The goal of this research is not to create or validate an instrument, but rather to use the most accepted one available. Therefore, the MUG-based model is used for this research. A discussion of the five MUG guidelines and their usefulness follows.

Content

Content reflects quality, completeness, and reliability of information included in a Web site
Does Web Usability Correlate with Usage?  Journal of International Technology and Information Management

(Venkatesh & Agarwal, 2006). The MUG guidelines suggest multiple subcategories that, in fact, capture various aspects associated with content. These subcategories are: relevance, relating to the pertinence of the content to the core audience; media use, signifying the appropriate use of multimedia content; depth/breadth, examining the appropriate range and detail of topics; and current and timely information, capturing the extent to which a Web site’s content is current. As Green and Pearson (2009) explain, “The content construct [is] similar to the technology acceptance constructs of perceived usefulness and relative advantage” (Agarwal & Prasad, 1997; Moore and Benbasat, 1991).

Ease of Use

Ease of use ties to an individual’s assessment of the mental effort involved in using a Web site (Venkatesh & Agarwal, 2006). The construct of ease of use has been employed extensively in information systems research (Venkatesh, Morris, Davis, & Davis, 2003) and has been shown to be an important predictor of technology acceptance outcomes. Well designed navigation and search options reflect the support provided to the user when moving in and around the site.

Promotion

Promotion captures the advertising of a Web site on the Internet and other media. Promotion should communicate the primary features, goals, or themes of the site. It has to convey an appealing attitude to its target audience (Keeker, 1997). Many Web sites failed to use their promotional spots to let people know that they have time-sensitive events and constantly fresh content (Venkatesh & Ramesh, 2006).

Made-for-the-medium

As explained by Green and Pearson (2009), Made-for-the-medium relates to fitting a Web site to a particular user’s needs. Peppers and Rogers (1999) suggest contemporary marketing strategies require Web sites with dynamic content tailored to specific user needs.

Emotion


These five dimensions have cited repeatedly, and at the time of data collection their aggregation into MUG is the most accepted usability instrument in the information systems literature.

An important question is whether the usability of a Web site correlates with usage of that Web site, because if it does then companies would logically want to pay increasing attention to the site’s usability. This study attempts to answer that question, by comparing the usability rating of a site with its actual usage. As mentioned, usability will be measured with MUG, and usage
data, including Unique Audience, Time Per Person, and Total Web Page Views, has been provided by Nielsen Online (http://www.nielsen-online.com).

RESEARCH METHODOLOGY AND DATA COLLECTION

This research used eight different retail Web sites from the Fortune 500 to study consumer-focused Web site usability and corresponding usage. Table 1 lists the companies and associated Web sites used, as well as the number of survey respondents per Web site. The study consisted of the review of these commercial retail Web sites by students majoring in Business who served as customer surrogates to respond to survey questions concerning Web site usability. University students are considered an appropriate surrogate in consumer research that examines products the students are likely to purchase on a regular basis (Lynch, Calder, Phillips, & Tybout, 1982). The survey asked multiple questions about each of the five areas of usability outlined in MUG, as well as four general questions about usability. The intent of this research was to use a previously validated measure of usability, and therefore the MUG portion of the survey followed the guidelines outlined by Venkatesh and Ramesh (2006), including the weights from their United States survey (Content = 32, Ease of Use = 12, Promotion = 9, Made-for-the-Medium = 37, and Emotion = 10) to calculate an overall usability rating. Survey questions used for this study are listed in the Appendix.

Each Web site was visited once after November 2007 and again after January 2008. The students received a small amount of class participation extra credit to complete the survey. A short paragraph describing the Web site which each student visited was also required, in an attempt to verify that the respondent was going to the correct site and answering the questions carefully. Web sites were assigned randomly to those participating, and sites and surveys were given to large lecture classes of “Information Systems in Organizations”, the required information systems class in the College of Business. Major and year demographics were not asked in the survey, but the surveys were completed each semester by student groups which were approximately 65% Juniors and 35%, with major percentages being nearly identical each semester as follows: Marketing 30%, Finance 25%, Business Administration 20%, Human Resource Management 8%, Organizational Management 8%, Operations and Information Management 6%, Undeclared Business 3%. Two hundred and twelve students were given the opportunity to participate in the November study, and 151 took the opportunity, for a response rate of 71%. One hundred and seventy-four different students were given the opportunity to participate in the January study, and 101 took the opportunity, for a response rate of 58%. All measures in the questionnaire were done on a seven-point Likert scale ranging from “completely disagree” (1) to “completely agree” (7).
Table 1: Web Sites Visited and Number of Survey Responses per Site.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.dillards.com">www.dillards.com</a></td>
<td>Dillards</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td><a href="http://www.jcpenney.net">www.jcpenney.net</a></td>
<td>JC Penny</td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td><a href="http://www.kohls.com">www.kohls.com</a></td>
<td>Kohls</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td><a href="http://www.nordstrom.com">www.nordstrom.com</a></td>
<td>Nordstrom</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td><a href="http://www.saksincorporated.com">www.saksincorporated.com</a></td>
<td>Saks</td>
<td>18</td>
<td>14</td>
</tr>
<tr>
<td><a href="http://www.searsholdings.com">www.searsholdings.com</a></td>
<td>Sears Holdings</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td><a href="http://www.target.com">www.target.com</a></td>
<td>Target</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td><a href="http://www.walmart.com">www.walmart.com</a></td>
<td>Wal-Mart</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td></td>
<td><strong>151</strong></td>
<td><strong>101</strong></td>
</tr>
</tbody>
</table>

Survey results were then compared with Web site usage data from Nielsen Online to compare usability numbers with usage numbers. Acknowledging that many factors can influence a Web site’s overall traffic, notably the holiday retail season of late fall, this research only considered relative increases and decreases in Time (Seconds) Per Person and Page Views Per Person, as compared to relative increases and decreases in rated usability. December data was omitted, although this comparison method should be valid even during the high holiday season. In other words, if a retail site experienced unusually heavy volume in November (or December) and much less in January, Time (Seconds) Per Person and Page Views Per Person should still relate to usability. This research considered per person traffic rather than overall traffic. Similarly, the companies compared have, in some cases, vastly different online market shares, partly due to brand name recognition and advertising, and this comparison method allows for that. If one site received one thousand visitors in November and January, and a different site received one million, looking at per person time and per person page views on the site “levels the playing field” and allows for a meaningful usability comparison. Finally, this study is not attempting to rate one site as more usable than another. Although the ratings could be viewed in that regard, the goal is to examine changes in usability and usage over time.

RESULTS

Survey results as well as Web site usage data is shown in Table 2. “Average Usability” is the average of the four survey questions on usability. “Average Calculated Usability” is the average of the calculated usability ratings, using the survey questions covering the five areas in the MUG as well as the weighting from Venkatesh and Ramesh (2006). “Seconds / Visitor” is the number of seconds each visitor stayed on the Web site, and “Page Views / Visitor” is the number of total page views divided by the number of unique visitors to that site for the month. As stated earlier, usage data was provided by Nielsen Online. All four of these statistics are shown for both November 2007 and January 2008.
Table 2: Usability and Web Site Usage.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dillards</td>
<td>4.93</td>
<td>5.26</td>
<td>481</td>
<td>19</td>
<td>5.14</td>
<td>5.41</td>
</tr>
<tr>
<td>JC Penny</td>
<td>4.40</td>
<td>4.67</td>
<td>1021</td>
<td>42</td>
<td>4.21</td>
<td>4.67</td>
</tr>
<tr>
<td>Kohls</td>
<td>5.50</td>
<td>5.32</td>
<td>782</td>
<td>34</td>
<td>5.17</td>
<td>5.16</td>
</tr>
<tr>
<td>Nordstrom</td>
<td>5.17</td>
<td>5.37</td>
<td>505</td>
<td>20</td>
<td>5.23</td>
<td>5.56</td>
</tr>
<tr>
<td>Saks</td>
<td>3.83</td>
<td>4.16</td>
<td>249</td>
<td>18</td>
<td>3.82</td>
<td>4.29</td>
</tr>
<tr>
<td>Sears Holdings</td>
<td>3.85</td>
<td>4.20</td>
<td>585</td>
<td>18</td>
<td>3.72</td>
<td>4.40</td>
</tr>
<tr>
<td>Target</td>
<td>5.53</td>
<td>5.37</td>
<td>603</td>
<td>20</td>
<td>4.63</td>
<td>4.74</td>
</tr>
<tr>
<td>Wal-Mart</td>
<td>5.75</td>
<td>5.52</td>
<td>919</td>
<td>28</td>
<td>5.02</td>
<td>5.08</td>
</tr>
</tbody>
</table>

Table 3 shows the changes in each of the four statistics moving from November to January, as well as the significance for usability at $\alpha = .15$. For example, the “Average Usability” for Dillards was rated 4.93 for November and 5.14 for January, thus an increase of 0.21. The “Calculated Usability” for Dillards was rated 5.26 for November and 5.41 for January, thus an increase of 0.15, and the increase of the aggregation of average and calculated usability was significant at $p < 0.140$.

Table 3: Usability and Web Site Usage Changes from November to January.
Lastly, Table 4 shows the Correlation Matrix with correlations between the four statistics.

Table 4: Correlation Matrix.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Change: Survey Avg. Usability</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change: Survey Avg. Calculated Usability</td>
<td>0.9558</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change: Seconds / Visitor</td>
<td>0.9247</td>
<td>0.8851</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>Change: Page Views / Visitor</td>
<td>0.9107</td>
<td>0.8828</td>
<td>0.7690</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

DISCUSSION AND CONCLUSIONS

The results are extremely promising. The correlation coefficients ("\( r \)\) in the Correlation Matrix are very high. A common rule of thumb is that 0.7 or higher indicates a strong correlation between two sets of data. We would expect that Usability and Calculated Usability would move in the same direction (\( \alpha = .15 \)), as well as Seconds and Page Views (\( \alpha = .7690 \)). What’s more important is that both of the usability ratings moved in the same direction as both of the usage statistics. Usability was correlated with Seconds at .9247 and with Page Views at .9107. Calculated Usability was correlated with Seconds at .8851 and with Page Views at .8828.

On a more detailed level, three Web sites had both their Usability and Calculated Usability go down (\( \alpha = .15 \)) from November to January (Kohl’s at \( p < 0.129 \), Target at \( p < 0.041 \), and Walmart at \( p < 0.071 \)), and Seconds per Visitor and Page Views also went down. Two Web sites had their usability ratings go up (\( \alpha = .15 \)) from November to January (Dillards at \( p < 0.140 \) and Nordstrom at \( p < 0.149 \)), and their Seconds per Visitor and Page Views also went up. The other three Web sites, JC Penny’s, Saks, and Sears Holdings, had mixed results closer to no change from November to January (not statistically significant at \( \alpha = .15 \)), and this was reflected in the usage statistics as well. The high correlations between the changes of all sites indicate that changes in usability calculations from a survey strongly indicate changes in usage.

As stated previously, the goal of this research is not to rate Web sites as “good”, or even “usable”. Rather, it’s to demonstrate that the academic research to date on usability is quite helpful in determining relative direction and magnitude of a site’s usability over time, and that could translate into proportionally more or less usage and site “stickiness”. Therefore, calculating usability might be wise for business managers monitoring or contemplating Web site design changes, promotions, etc.
RESEARCH LIMITATIONS

Usability studies are continuing, and new dimensions may emerge to gain a more robust measurement of usability. As these studies expand, additional constructs may be added to MUG to produce a modified or new instrument. This study used MUG, the most valid instrument at the time of data collection, to conduct the research. If constructs change, new usage data will need to be collected to validate the results of this study.

Additionally, there are three sample concerns. First, the November study (later in the semester) had a response rate of 71% and the January study (earlier in the semester) had a response rate of 58%. One possible hypothesis for this difference would be that students became more concerned with their grades later in the semester (as students often do) and were more motivated to participate. We are confident that this did not distort the results, because once a student did participate he/she was required to complete a paragraph about the site and complete the entire survey. So even though the response rate went down, we do not believe the quality of the responses themselves went down. However, to be certain future studies should have students complete the survey at the exact same time during the semester. Second, we had no way of knowing for sure how long the students stayed on the actual web site and how thoroughly they reviewed its usability. While we feel comfortable that the descriptions received from the students describing their web site visits do indicate serious site examinations, future studies would be wise to time the students on their site visits, or, most preferably, have the students visit the sites in a controlled lab environment. And third, it would be desirable to have a data sample where the usability went up and/or down more dramatically from period to period so the correlation with usage would gain augmented credibility.

And finally, while highly promising, it should be noted that the results are spanning eight Web sites and using only simple statistics (averages and correlation) with a large α (.15). Additional studies comparing more sites with more rigorous statistics are needed, and these results provide good reason to undertake such efforts.

REFERENCES


**APPENDIX**

**SURVEY INSTRUMENT – MUG QUESTIONS USED**

**Web Site Usability Survey** (modified to show only questions used in this study)

Please type the site address you have just visited:

In 150 words or less, describe the appearance of the site and your experience of the site visit:

Please rate the following by checking the number that best reflects your opinion of the Web site you just visited:

1 - Strongly Disagree
2 - Disagree
3 - Slightly Disagree
4 - Neither Agree or Disagree
5 - Slightly Agree
6 - Agree
7 - Strongly Agree

**CONTENT**

I feel this Web site provided information relevant to the customer.

Strongly Disagree  1  2  3  4  5  6  7  Strongly Agree

I feel this Web site offered personalized information and layout.

Strongly Disagree  1  2  3  4  5  6  7  Strongly Agree
I feel this Web site provided timely information.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

I feel the amount of information displayed on the Web site was adequate.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

EASE OF USE

I find it easy to get this Web site to do what I want it to do.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

The sequencing and navigation on this Web site were clear.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

The layout of pages made tasks easier.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

The search functions provided in this Web site helped me find relevant information.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

PROMOTION

I feel this Web site provided a good promotion for the products/services of the company.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

I feel this Web site I am browsing is promoted well externally on other Web sites and/or other media.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

MADE-FOR-THE-MEDIUM

I feel engaged/involved by the interactivity of the site.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

The extent to which this Web site can be tailored to fit my specific needs was adequate.

Strongly Disagree   1  2  3  4  5  6  7  Strongly Agree

I feel that this Web site provided me the opportunity to be part of an online group or community.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

I feel this Web site reflects most current trend(s) and provides nice design for the site visit.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

**EMOTION**

I feel this Web site provided features to promote customers’ excitement.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

The visit of this Web site was enjoyable.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

**OVERALL USABILITY**

I would be willing to visit this Web site again.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

I would be willing to recommend this Web site to others.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

I would be willing to purchase from this Web site if needed.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

I have positive things to say about this Web site.

Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Gender:

Male  Female

To what age group do you belong?

17-20  21-25  26-30  31-35  36-40  41-45  46-50  Over 50

What statement best describe your level of experience of using Web?

I have used the Web a few times before this survey

I use the Web a few time a month

I use the Web every week

I use the Web almost every day

Thank you for taking this survey! Please Click the Submit Button.

Submit