

Increased popularity through compliance with usability guidelines in e-learning Web sites

Journal Article**Author(s):**

Scowen, Greg; Regenbrecht, Holger

Publication date:

2009

Permanent link:

<https://doi.org/10.3929/ethz-a-005921621>

Rights / license:

[In Copyright - Non-Commercial Use Permitted](#)

Originally published in:

International Journal of Information Technology and Web Engineering 4(3), <https://doi.org/10.4018/jitwe.2009100603>

INTERNATIONAL JOURNAL OF INFORMATION TECHNOLOGY AND WEB ENGINEERING

July-September 2009, Vol. 4, No. 3

Table of Contents

RESEARCH ARTICLES

- 1 **SWAMI:
A Multiagent, Active Representation of a User's Browsing Interests**
Mark Kilfoil, University of New Brunswick, Canada
Ali Ghorbani, University of New Brunswick, Canada

- 25 **Analyzing the Long Range Dependence and Object Popularity in Evaluating the Performance of Web Caching**
G. P. Sajeev, Government Engineering College, Kozhikode, India
M. P. Sebastian, National Institute of Technology, Calicut, India

- 38 **Increased Popularity Through Compliance with Usability Guidelines in E-Learning Web Sites**
Greg Scowen, ETH Zurich, Switzerland
Holger Regenbrecht, University of Otago, Dunedin, New Zealand

- 58 **Design of an Integrated Web Services Brokering System**
Frederick Petry, Naval Research Laboratory, Stennis Space Center, USA
Roy Ladner, Naval Meteorology and Oceanography Command, Stennis Space Center, USA
Kalyan Moy Gupta, Knexus Research, USA
Philip Moore, Knexus Research, USA
David W. Aha, Naval Research Laboratory, USA
Bruce Lin, Naval Research Laboratory, Stennis Space Center, USA
Richard Sween, Naval Research Laboratory, Stennis Space Center, USA

BOOK REVIEW

- 78 **Web Engineering:
A Practitioner's Approach**
*Reviewed by Hussein Al-Bahadili, The Arab Academy for Banking and Financial Sciences,
Jordan*

Increased Popularity Through Compliance with Usability Guidelines in E-Learning Web Sites

Greg Scowen, ETH Zurich, Switzerland

Holger Regenbrecht, University of Otago, Dunedin, New Zealand

ABSTRACT

This article identifies and measures correlations between compliance with usability guidelines and the popularity of a Web site. A sample of e-learning Web sites was reviewed and their usability scored using a Web-based evaluation system developed during the study. This usability score was then tested against five different ranking systems using Spearman's Rank correlation. The results of these tests show a strong correlation between compliance with usability guidelines and Web site popularity. The five ranking systems also showed positive correlations to each-other and to the usability of the sites. The conclusion drawn from these results is that compliance with usability guidelines could be a way to achieve higher Web site popularity and visitor numbers.

Keywords: E-Learning, Electronic Learning, Usability, Web-Based Learning, Web Site Design, Web Site Development

INTRODUCTION

We have often suggested to clients and students that usability is an essential part of any Web site development and that compliance to usability guidelines could be a pathway towards a more popular site. Yet many times these suggestions have been ignored, or in the least respondents have not considered usability studies to be important to their project. Due to

a lack of previous research which examines the relationship between usability and popularity, convincing developers to invest resources into usability studies has been difficult. This research aims to answer the question: is there a positive relationship between compliance with usability guidelines and Web site popularity, and if so how strong and relevant is this relationship? Other questions addressed consider what usability guidelines should be adopted and how popularity can best be measured. The findings of this study are intended to create a foundation for further investigation into the effect of

DOI: 10.4018/jitwe.2009100603

compliance with usability guidelines on Web site popularity.

The scope and objectives of this research are:

- To identify academically established—industry recognized Web site usability guidelines.
- To measure the compliance with these guidelines in e-learning Web sites. The population of this study has been narrowed down from all Web sites to e-learning Web sites, with the intention that the representation of the findings in one industry is transferable to others.
- To measure the correlation between compliance with usability guidelines and the popularity of a Web site. An assumption has been made that members of a group of Web sites within a particular educational subject area, that is, English or Math, have the same chance as each other of becoming popular.
- To identify and measure correlations between compliance with usability guidelines and five Web site ranking systems.

The first section of this article reviews previous literature surrounding the fields of usability and Web site popularity. The method section follows on from the literature review and provides a detailed description of how this research was conducted. This is followed by the results and conclusions sections which include a discussion about what the findings could signify and what work still needs to be done.

PREVIOUS USABILITY AND POPULARITY RESEARCH

Overview

Tools and methods that may be used to prove a correlation between usability and Web site popularity have been reviewed. The tools concerned include established Web site design usability guidelines and software for testing

compliance to such guidelines. Suggested methods, as an alternative to tools, for testing compliance are also discussed and accompanied with a review of methods for measuring Web site popularity.

Usability

What is Usability?

Usability is a well established concept and is precisely defined by the widely accepted ISO9241 standard (Petrie & Kheir, 2007). Part 11 of ISO9241 defines usability as the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments (ISO, 1998). In this context *effectiveness* refers to the accuracy and completeness of the tasks, *efficiency* considers the resources expended in completing the tasks, and satisfaction means the comfort and acceptability of the system to its users. Prior to the development and wide-spread acceptance of ISO9241, perhaps the most commonly cited definition of usability was that of Jakob Nielsen who broke usability into five areas: learnability, efficiency, memorability, errors, and satisfaction (1993).

It is important, for the purpose of this study, to clarify the separation between usability and accessibility. While there are many similarities between usability and accessibility guidelines, the two fields are not the same. However, accessibility could be considered a subset of usability, since usability implies accessibility (Brajnik, 2000). If a Web site is usable by all users then it must be accessible. Brajnik also shows that the contrary is not necessarily true, that is, problems that may affect the usability of a page may not affect the accessibility. Likewise, a article produced in partnership with the National Cancer Institute makes an attempt, as the title suggests, at “Bridging the Gap: Between Accessibility and Usability” (Theofanos & Redish, 2003). Further reinforcing the subset theory, Theofanos and Redish note “meeting the required accessibility standards does not,

however, necessarily mean that a Web site is usable for people with disabilities” (p. 38).

Established Web Site Usability Guidelines

The definitions for usability by ISO (1998) and Nielsen (1993) are also suitable for Web sites, in that a Web site is a form of system. Defining suitable guidelines, however, with which to achieve this usability is more difficult. Attempting to further define Web site usability and appropriate guidelines are subjects of ongoing research by a number of Web site usability organizations. This research and less formal sources can be easily found through common search engines. However, the quantity of material available only compounds the problem because much of the information is conflicting. In a 1997 report where he attempted to create a resource guide for Web site usability, Ohnemus (1997) produced a list of 8 guide books, 22 Web style guides, 8 accessibility guides, and a further 10 resources on usability. Furthermore, the World Wide Web Consortium (or W3C) also produced a set of guidelines that quickly gained acceptance. These guidelines were officially focused on accessibility, although the confusion of whether they were usability guidelines was even prevalent among the W3C members responsible for them (W3C, 2001).

The conflict between all of the available guidelines did not go unnoticed and in 2001, the National Cancer Institute started working on identifying research related to Web design and usability. Their findings became the subject of panels at conferences (Koyani & Allison, 2003) and later developed into a set of 187 peer reviewed guidelines that are now published by the American Government for use in government Web sites (HHS, 2006). The effort made and resulting guidelines have been noted and praised:

“Prescriptive guidance is often voluminous, vague, conflicting, or divorced from the context in which sites are being developed, thus making it difficult to apply. Many guidelines have not been validated empirically and there is

little overlap across guideline sets. A noteworthy exception is the research-based guidelines that researchers at the National Cancer Institute developed” (Ivory & Megraw, 2005).

Due to the general support and the nature of the continual redevelopment/refinement by the HHS, this study will be based on a selection of their guidelines. The next question to answer is how best to test for compliance to these guidelines.

Testing a Web Site’s Compliance to the Guidelines

Normally a usability evaluation would be completed using one or both of two kinds of methods: usability inspection methods or user testing (Brajnik, 2000). An example of usability evaluation is Nielsen’s heuristic evaluation (Nielsen & Mack, 1994) where a group of experts use a Web site and assign severity levels to usability issues. By contrast, a normally non-specialist and randomly sampled group of system users conduct user testing. These users are given instructions of tasks to complete on the Web site and are asked to note system functions that do not work or that they do not like (Instone, 1999). Both of these methods confirm the findings of Preece et al. that “evaluation is concerned with gathering data about the usability of a design or product by a specified group of users for a particular activity within a specified environment or work context” (Preece et al. cited in Spiliopoulou, 2000).

However, for this study we do not need to fully evaluate the usability of a Web site to show that the chosen guidelines have been complied with. To conduct either usability inspections or user testing for a large number of Web sites would also be beyond the scope of this study due to time restraints and the quantity of information required. Rather, each Web site will be tested against a checklist of items derived from the chosen guidelines. Such a checklist can be tested through the use of an automated tool or by manual observation.

Brajnik (2000) compiled a comprehensive table for the comparison of evaluation

tool features and one tool, 'LIFT' (Usablenet, n.d.), stood out clearly above the rest, due to its coverage of more aspects of usability and extended features. When supplied a URL, the LIFT tool retrieves the text form of the Web site and analyses it according to various rules. These include, for example, checking for completeness of "Alt Tags" in images, use of Web safe colors, and the validity of HTML code. Unfortunately, the features measured by the LIFT tool were often inconsistent with the guidelines produced by the HHS. In fact, the majority of the items checked by automated tools seem to relate more to accessibility than usability as a whole. As previously discussed, merely meeting accessibility standards does not infer meeting usability standards. Wattenberg (2004) refers to LIFT and other tools from Brajnik's table as "automatic validation tools ... to help evaluate the accessibility levels of a completed Web site" (p. 14). He goes on to note that "these tools have also been found to bypass emerging technologies and miss important usability problems." Even when we only consider the automated tools as being accessibility measurement systems, Wattenberg points out that they cannot be considered efficient by discussing a study from Scotland where the potential to develop a single evaluative tool to help create accessible Web sites was conducted. "The researchers did not find a single tool or process that would achieve these goals" (Wattenberg, 2004, p. 14).

With automated testing declared unsuitable, the remaining option for testing compliance to usability guidelines is to manually mark items off on a checklist. The use of this method has been validated by Keevil (1998) who defines a checklist as "a list of questions that require a yes or no answer" (p. 271), although he also offers an alternative of using scoring instead of yes/no options (p. 273). Preferring the "Yes or No" concept, Keevil goes on to explain how to create a checklist. His suggestion is to select measurable attributes that are based on usability research and then to write these down in a question format (p. 274). An example of a question that may be included on a checklist could be: "Is some form of notice given (e.g.,

breadcrumbs) to identify where on the site the user is?" In his summary, Keevil provides a brief list of advantages and disadvantages to using a checklist system. One noted advantage is that a "checklist is inexpensive and easy to implement" (p. 275). The major disadvantage is that a bias may be introduced by evaluator's interpretations of the guidelines. However, according to Keevil, this bias is overcome through the use of a "Yes or No" technique. Keevil's work has been a major contributor to the design of this study. For the evaluation of compliance to usability guidelines, a checklist with the "Yes or No" design will be used, with the addition of a 50% option. Where a guideline has been adhered to in part, but perhaps not in full or throughout all assessed parts of the system, then a 50% or "partially complied" grade will be given. Further to Keevil's support for the checklist evaluation system, Human Factors International—a large and experienced user-centered design group—have also touted the use of a similar checklist system called "Usability Scorecards" (Weinschenk, 2007).

Web Site Popularity

What Is Web Site Popularity?

Unlike Web site usability, the definition of Web site popularity is not so clear. First, it is important to discuss the relationship between success and popularity. If popularity is taken to mean the state of being "liked or admired by many or by a particular group" (Oxford, n.d.) then success is not necessarily equal to popularity. Belanger et al. (2006) note that since success is goal-specific, it is defined differently depending on the needs of the business. For an online retailer, success may mean the percentage of visitors converted to buyers, or the degree of loyalty as shown by the number of return visitors (Schonberg, Cofino, Hoch, Podlaseck, & Spraragen, 2000). The quantity of site traffic is, however, still recognized as "... the predominant way of determining success from organizations' perspective" (Belanger et al., 2006). For the purposes of this study, we have defined Web

site popularity based on the Oxford definition of popularity above: A popular Web site is one that is liked or admired by many people.

How a Web site becomes popular is the topic of much discussion. Adamic and Huberman (2001) report that Web site growth follows power laws, allowing for the growth of a Web site to be mathematically predictable: "The day to day fluctuations in the number of visitors to a site is proportional to the number of visitors the site receives on an average day" (p. 58). The visitors to a site are of two types; repeat visitors who in turn influence new visitors, and new visitors who have come through a referral or advertising. The more visitors a site has, the more referrals they get and the more advertising they can afford. This idea of growth infers that popular pages will always become more popular than less popular pages. Smaller Web sites do not have the same chance to grow. This theory is supported by Cho and Roy (2004) who report that popularity-based search engine rankings are biased against unknown pages: "When search engines constantly return popular pages at the top of their search results, more Web users will 'discover' and look at those pages, increasing their popularity even further" (p. 20-21). In this case, popular pages are those that have high visitor numbers, rather than inbound links as are used in this study. This finding is also echoed by Kavassalis, Lelis, Rafea, and Hardi who state that "... users are thus more likely to learn about popular pages than unpopular ones" (2004). The question should be asked, however, of how sites that are newer can have come to be much more popular than older sites. The power law theory of popularity does not take into account how a popular site became popular in the first place. An example to consider would be that of Google vs. Yahoo. For many years, Yahoo was a search engine that was widely known and used. Suddenly, in 1998, Google launched a similar service, albeit with a new approach and differing methods behind the system. Within a few years, Google had become the search engine of choice for more Internet users than the others. What caused this? Could it be that Google was more usable?

Another report from De Angeli, Sutcliffe, and Hartman (2006) explains how they conducted user testing on two Web sites with identical content but different navigational systems and layouts. One site had a traditional menu-driven layout; the other was a metaphor-based interactive and animated design. Their findings clearly showed that the different designs had more or less appeal and usability levels dependent on the age group of the users. Therefore, we also need to question what influence the age of Internet users has on Web site popularity. If the majority of Internet users fall into a certain age group, will a particular style of Web site also have an increased potential to become popular? With this question in mind, this study will use a sample of Web sites from the educational sector with the aim of limiting the potential age differential in users as much as possible.

Measuring Web Site Popularity

There are two common methods for measuring Web site popularity. One of these is to use a statistical program to count the number of visitors to the site and compare this with other sites. The alternative method mirrors the academic world (and as shown below the concept of Google PageRank) and is measured by counting the number of inbound links to a Web site, which are considered as referrals in support of a Web site.

Attempting to count the number of visitors to all Web sites is an impossible task. To do this, every Web site host in the whole world would need to submit their user data to one central location. This could be achieved through the use of automated software or by manual means, but overseeing such a task would require huge infrastructure and a lot of financial resources. This does not even take into consideration the ethical and business reasons that would stop a large portion of these hosts from sharing that information. This conundrum has resulted in various companies establishing services or tools that count traffic of competing Web sites. An example of a well-known traffic counting service is the Internet media and market

research company The Nielsen Company and their "Nielsen Netratings" (n.d.). This company charges clients a fee for providing them with traffic statistics for their Web sites and those of their immediate competitors. Unfortunately, the fees involved are well beyond the budget afforded to this study. An alternative tool is available for free, that of Alexa.com (Alexa, 2007), a subsidiary of the Amazon Company that provides a downloadable tool-bar for Internet Explorer users. This toolbar offers search engine functionality and information about the sites that users are browsing. In the background it also records traffic information and saves this data to their Web site, which in turn reports the traffic rating of Web sites. Although this information sounds of promise to this study it not altogether helpful. This is because "Alexa's sample is known to be biased towards users of Microsoft Windows, particularly those who use Microsoft Internet Explorer" (Sullivan & Matson, 2000, p. 141). Since users of other operating systems or browsers are not recorded, and traffic from all other Internet users is not counted, the resulting traffic statistics are not very helpful at all. With this limited scope in mind, the Alexa rating for sites will be reported in this study, but will not be the primary indicator of popularity.

Conversely, the concept of links toward a Web site being positive referrals to a site makes a lot of sense. Even when a Web site links to a site the owners do not necessarily like, they are still suggesting that other people look at it, so therefore it is a positive referral in at least one sense. Brin and Page understood this concept when they designed the search engine Google (1998). They created a system called PageRank, where links pointing to a page are accorded a quality status (depending on the pages they appear on) and counted. A page, therefore, has a calculated PageRank that is based on the number of links pointing to it and the quality of those links. Through PageRank, the position of a page in Google search engine results is decided. As with the Alexa rating, the PageRank of each site's main page will be recorded, however, for this study the popularity rating of each Web site will be assessed through the number

of links (treated as referrals) located through three sources: Google, Yahoo, and Del.icio.us. Because Google's PageRank system puts Web sites and links through tough checks and removes spam content, the number of links in Google (Google link-count) will be the chief indicator of popularity in this study.

Yahoo and Del.icio.us have been selected as additional link-count sources due to their popularity and the availability of their statistics. Yahoo is a search engine that operates in a similar fashion to Google and will be used to reinforce the Google results. Del.icio.us (2007), on the other hand, considers the social aspect of the Web. The Del.icio.us site allows for Web users to create bookmark/favorite lists publicly, thus allowing for other users to discover sites based on what their peers like. Unlike the process used by Google's PageRank system, in this study the quality of the sources of links will not be considered.

A Proposed Correlation Between Usability and Web Site Popularity

The objective of this study is to show that a correlation exists between adherence to usability best practices and the popularity of a Web site. There are a few studies that propose a correlation between usability and popularity; however, none of these actually measure the existence of a relationship between the two or set out to evaluate this relationship.

One article that suggests that usability is important to the popularity of a Web site contains comments that appear, unfortunately, to be based on the author's opinion rather than evidenced fact. In his article "User Interface Directions for the Web," Jakob Nielsen brashly states:

"Unfortunately, it is common for sites to aim at being 'cool', 'sizzling', or even 'killers' rather than trying to do anything for their users.... Design Darwinism will tend to drive out the most flamboyant sites and concentrate traffic at sites that follow the usability principles" (Nielsen, 1999).

It is clear that when he made this statement, Nielsen had not taken into account sites that

contain numerous usability issues yet remain largely popular. A modern-day example can be found at MySpace.com and when Nielsen wrote his report there were similar examples, including the MySpace of the 90's, geocities.com. There will always be exceptional Web sites that become popular regardless of their flaws. This is due to the overwhelming social drive behind the Web. As discussed by Malcolm Gladwell, there can become a point in the life-cycle of a product, service, or system where it can simply cross a threshold, tip, and spread like wildfire (2000). Other comments regarding the importance of usability to Web site growth seem to carry a little more credibility. Weinschenk reminds developers, marketers, and technology managers that the key to success remains that a product or service is actually useful and usable (Weinschenk, 2007). This supports the theory that usable Web sites will be more popular than unusable Web sites. Lederer, Maupin, Sena, and Zhuang confirm that "use of Web sites is to some extent dependent on the usefulness of the information content and ease of using the site" (1998, p. 200). They go on to suggest that to encourage visitors to their sites, Web managers should focus on usability (p. 201).

A report by Sullivan and Matson, "Barriers to Use: Usability and Content Accessibility on the Web's Most Popular Sites" (2000) is one of the key resources used in the preparation of this study. Sullivan and Matson made an analysis of the Web's 50 most highly trafficked sites, as identified through the use of the previously discussed Alexa toolbar (p. 141). While the study claims to have assessed the usability of their sample sites, the tool used for this assessment was LIFT (p. 142), the same tool discussed above. As has been previously established though this review of literature, LIFT does not adequately assess usability, rather it focuses on accessibility issues. Therefore the results of this report pertaining to the usability of the sites reviewed are not particularly reliable. Sullivan and Matson have pointed out the bias of Alexa, but failed to acknowledge the unsuitability of the LIFT tool to usability analysis. Moreover, they claim that their findings "...suggest that a

meaningful ordinal ranking of content accessibility... correlates significantly with the results of independent automated usability assessment procedures" (p. 139) based on the use of the LIFT tool. The relevance of their findings is diminished when the LIFT tool is inspected and shown to really assess accessibility rather than usability. Instead of assessing both usability and accessibility, Sullivan and Matson appear to have assessed accessibility twice with differing tools and have then shown a correlation between the consequential results. However, because accessibility is a subset of usability, it is true that when a site is made more accessible it is inevitably also made more usable to at least some users.

There are, however, some insightful remarks to be taken from Sullivan and Matson's work that have been considered in the design of this study. These include the review of the Alexa toolbar which has been used to help design what does and does not receive consideration for ranking popularity (p. 141) and the concept of treating the homepage as representative of the whole site (p. 141). They also conclude that although guidelines are widely available, publicized, and known about they appear to be largely ignored. This serves as confirmation of the researcher's beliefs and as motivation for this study. If it can be shown, as predicted, that compliance with usability guidelines correlates with Web site popularity perhaps more organizations will invest time and effort into this field.

Summary of Literature Review

Reviewing literature related to this study has answered many questions, but also created new questions that need answering. A gap in knowledge has been identified. Does compliance with usability guidelines correlate to increased Web site popularity? It is obvious that some researchers are hinting at this but there does not appear to be any previous study that attempts to answer that question.

The review has also helped to identify appropriate tools and methods for conducting

such a study, while identifying others that are not so suitable. Through reading the research of others a clear definition of usability has been established and the definition of popularity has been refined and stated, in the least for how it is to be used for the purpose of this study.

Key contributors to the field have been identified and their findings taken into consideration. With this solid foundation to build upon, we conducted this study.

METHOD

The Sample—Web Sites for Review

How Sites Originally Identified

We have defined e-learning Web sites as sites that contain educational resources such as self-testing software, games, templates, lesson plans, and tutorials. This study will further minimize the population by including only sites that contain resources for the K-12 sector of the education industry (hereafter referred to as Educational Resource Sites). Dr. Keryn Pratt, the tutor of a University of Otago course called “ICT in Education” (EDUX317, 2007) supplied the sample. One of the assignments that students of this course complete involves identifying three educational resource Web sites in a subject area of their choosing, and then reviewing these Web sites. Dr. Pratt agreed to supply a list of more than 200 of these identified Web sites from her files. The list included Web sites from various subject areas and did not include information about the original students or their reviews. From this list, it was intended to create groups by subject of no less than 10 sites. However, following a selection according to the criteria below, only 32 sites in total were included in the study and only two groups identified.

Selection Criteria for Sites

In attempt to minimize as many external influences as possible, official Web sites of universities, government organizations, and

schools were not included in the study. Because university Web sites are frequently visited by most students of that university, a larger institution would have a significantly higher visitor count. Government sponsored Web sites are often supported by large televised advertising campaigns, and furthermore some sites are a prescribed part of a curriculum, meaning that they are used regardless of their popularity or appeal.

This, essentially, leaves sites that are developed by independent companies or organizations and hobby groups to be considered. An example of a site that would be suitable could be a resource site developed by a nationwide astronomical club to help attract new members. The sites should all have content that is suitable for, or directed at, the K-12 sector of the education industry, meaning resources for primary and secondary education. All Web sites included in the study must have English as the main language of navigation and content to minimize the risk of the language of the intended audience acting as a compounding variable.

Other than this, the base criterion of selection was that the Web site was operational at the time of review.

Exceptions to Selection

One Web site that is provided by the University of Texas was included in the study. The “World Lecture Hall” site (<http://web.austin.utexas.edu/wlh/>) is a sub-site within the University’s main Web site and is an open resource for any interested visitors. The homepage used for data purposes and review was a second level directory in the main site, so no data was inherited from the University Web site itself.

There were no exceptions on Web sites that were not included in the study.

Problems Encountered in Site Selection

Having checked through the original supplied list of 225 sites supplied, only 38 sites met the criteria for inclusion in the study. One group

of 10 Web sites was identified that dealt with a diverse range of subject areas, and one other group of 7 Web sites that included resources for business education. The remaining sites formed groups of two or three in more individual subject areas such as physical education or biology.

Due to the original Web site list being created by New Zealand Teaching students, a few sites included in the study are on New Zealand domains. These Web sites are inherently likely to have lower visitor numbers than those on international domains. However, without including New Zealand and United Kingdom sites there would not be a sufficient quantity of sites in the study.

Measuring Usability and Popularity

Choosing Usability Guidelines

Initially we were prepared to collate a selection of numerous governmental and industry produced guideline documents, and to then compare these, charting the guidelines that were consistently agreed upon. During the collation process, a guideline document was encountered that had been produced by the US Department of Health and Human Services (HHS, 2006). This guideline document consists of 209 usability guidelines that have been collated through the collaborative effort of 18 academics and industry professionals. The result is a reliable, quantified and peer reviewed set of guidelines that do not exist anywhere else (HHS, 2006). The discovery of this document provides a sound base for this study.

We narrowed down the 209 guidelines to a more manageable selection of points that could be tested. The HHS guideline document applies “Strength of Evidence” and “Relative Importance” scores to each guideline which helped in this process. “Strength of Evidence” has been determined by a panel of eight usability researchers, practitioners and authors, and considers how well established each guideline is in academia and industry. “Relative Importance” refers to how important each guideline is to the success of a Web site. This

was determined by a panel of 16 reviewers, half of whom are usability experts, the others Web site developers. Because the testability of each guideline determined which ones could be used, this study included a broad range of both of these scores.

The first criterion for selecting a guideline for inclusion on this study was the ability to test conclusive compliance to the guideline visually by viewing a Web site. Subjective guidelines such as “Provide useful content” and development technique guidelines like “Establish user requirements” could therefore not be included.

A further criterion was the relative importance of a guideline. The HHS Guidelines “Relative Importance” scores fall between 1 and 5. For this study, only guidelines with a score of 3 or more were considered for selection. After identifying a set of 33 guidelines in the first selection process, a further five were removed according to the same criteria in a second round. This left a remaining set of 28 guidelines.

The Usability Score Concept

The selected guidelines had a range of relative importance scores associated with them, with 10 having a relative importance of 5, a further 16 with an importance of 4, and two with an importance of 3. From these relative importance rankings a ‘Usability Score’ was developed. If a Web site complied with all guidelines, the combined total of relative importance scores achieved would be 120. Assuming that compliance to more important guidelines makes a site more usable, we decided to measure the usability of a site based on this total, rather than based on a simple count of how many guidelines are complied. This ‘Usability Score’ was reported both as a total out of 120 and as a percentage.

Measures of Popularity

Five measurements of Web site popularity will be obtained during the data collection process. All of these will be reported in the results; how-

ever, the most importance will be placed on the number of links pointing to the site as found on the Google search engine. Further link-count measures included are sourced from Yahoo and Del.icio.us. Both of these are included to further support the Google link-count and to identify if there are similarities between these three sources. Additionally, the Google PageRank and Alexa rating will be reported to identify if these have a correlation with usability, although neither can be relied upon as an indicator of popularity.

Conducting the Reviews

Method of Reviewing Sites

Early in the study it was determined that individually reviewing each Web site and entering responses into a document would be very time consuming. Potentially, this would also result in discrepancies due to the difference in dates when the sites are reviewed and the mood of the reviewer. Automated options were considered, however these were found to be lacking in either their abilities for testing the guidelines or their reliability.

To combat these problems, the first author designed a Web-based evaluation tool specifically for use in this study. Using HTML frames, PHP scripting, and a MySQL database, the tool would enable the site under review to be presented alongside review questions and buttons for navigating the review process. As each question is answered, the response is saved to a database and the next question appears on screen with a refreshed copy of the review-site's homepage. Following the 28 guideline questions, five further questions asking for the rankings of Google, Del.icio.us, Yahoo, and Alexa would be presented.

Creating the Review Software (WES)

Web site Evaluation System (WES) was created using simple HTML framing techniques. A Frameset was created with a narrow band across the top of the page where each guideline could

be presented as a question, and the remainder of the browser window showed the site being reviewed. Adjacent to the question were four response options in the form of a radio-button group, and a button used for submitting the response and moving to the next question. All questions were written in a way that a "Yes" response equaled to compliance with the guidelines. A "No" response meant no compliance was visible. Furthermore, options were provided to indicate "partial compliance" with a guideline, or to "skip" the guideline if it wasn't applicable. Both of these options allocated a score of half the relative importance of the guideline considered. A guideline that might be skipped, for example, is "Label data entry fields consistently." If no forms were found on the site, this guideline was skipped and half points given to avoid inconsistencies in data.

The use of frames in this manner resulted that the site could be negotiated in the lower window of the browser without the question area being affected. This enabled the reviewer to locate instances of compliance to guidelines more efficiently. When a response was selected and the submit button clicked, the data was submitted via a script to a database and the next question would appear with a refreshed copy of the homepage of the review site presented. Because each response was directly submitted to the database, if a connection failure occurred, the review could be easily continued from the point of failure.

To assist in the retrieval of rankings from five different sources, five extra questions were presented following the 28 guideline questions. These questions asked for a textual input of the various ranks used in the study. To assist in answering these questions, a link was provided to a source for each ranking. These links took the URL of the site being reviewed as a variable and opened a new browser window where the required data was automatically shown. The same technique was used to obtain the download speed of sites for one of the guidelines.

Upon completion of all 33 questions, a "submit and review" button was provided. Upon clicking this, the scores were tallied, a

usability score as a percentage was calculated, and all review data was presented on the screen. These reports would be printed for each Web site reviewed. The results of all site reviews were stored in the MySQL database and then exported to an Excel spreadsheet, and later to SPSS.

Gathering Data

All 38 sites were reviewed over a two-day period in August 2007. The first author was the only reviewer involved in the study and treated all sites with the same criteria for grading the compliance to guidelines. The review was conducted using Internet Explorer version 7 on a Windows Vista operating system. The screen resolution was 1280 x 800 in wide-screen format on a laptop PC. Horizontal scrolling measurements were based on a 1024px wide resolution. Download times were calculated by an automated third-party system and were measured against a benchmark of downloading the homepage on 56k modem in 10 seconds or less.

The number of incoming links on Google and Yahoo were based only on links from external sites where possible. Both the Google and Yahoo link counting techniques used did not count links from within the base site itself.

Data Analysis

Breaking the Sites into Subject Categories

To create meaningful data the sites reviewed were broken down into categories based on subject matter. Web sites that deal with a diverse range of subjects are inherently going to receive more visitors than a site that includes resources on Astronomy for example. Sites that included more than five subject areas were allocated as members of the "Diverse" category.

Measuring the Correlation

To measure the strength of a correlation between compliance with usability guidelines and Web site popularity, all data obtained through the

review of sites was imported into SPSS for analysis.

The chosen method of measurement applied to this data was Spearman's Rank Correlation (Spearman, 1904). This method was chosen over the widely-known Pearson Product Moment method due to its suitability to non-linear, non-normal, ordinal data and also to ranked variables (McDonald, 2006). The number of links in Google, Yahoo, and Del.icio.us are non-normal in their distribution, and both Alexa and Google PageRank are ranks. The Alexa ranking system is a reverse order rank, so a correlation between the Alexa Rating and usability would be represented by a negative correlation coefficient.

Possible Factors that Skew Data

Some of the sites reviewed operated in a link-farm type manner, or were part of a network of sites that link to them exceedingly. A link-farm is a Web site that excessively trades links with other sites to artificially inflate their importance. Link-farms often contain little unique content and are more of a directory to other content.

Another external factor that may affect data is the location of the Web site, or its national centrality. Some sites reviewed are on the New Zealand domain (.co.nz) and are targeted at New Zealand students. Naturally, these sites don't attract as many visitors as a site on the .com domain which is international in its reach.

Subject area is also very relevant to the data. One site that scored very highly for usability was an Astronomy site. The popularity of the site was much lower than some other subject areas such as Internet, which is relative to the popularity of the subject itself. This is why the final research data is broken into groups based on subject.

RESULTS

Compliance to Usability Guidelines

An objective of this study was to measure the compliance with usability guidelines within a

group of e-learning Web sites. The results of the review of 28 selected guidelines over 38 Web sites is presented in Appendix Table 1.

Two of the tested usability guidelines were complied to most frequently: "Eliminate horizontal scrolling" and "use mixed-case for prose text." Both of these guidelines were complied with in 36 of the 38 sites reviewed. Horizontal scrolling was considered to be eliminated if no horizontal scrollbar was displayed when viewing the site at a resolution 1024 pixels wide. None of the reviewed Web sites completely failed to comply with the prose-text recommendations. The two sites that did not fully comply had some segments of prose-text that were capitalized, with the majority being correctly formatted.

The least frequently complied guideline was "Minimize page download times," which required that the homepage of the Web site being reviewed download in 10 seconds or less on a 56k modem. This was tested with the support of an external source; Only 5 of the 38 Web sites tested complied with the 10 second requirement. Because this guideline could only be complied with completely or not at all it was also the guideline most frequently not complied with, with 86.8% of sites failing to download in 10 seconds or less.

Of the 28 guidelines, the majority of these (23) were more often complied with than not. Four guidelines exhibited a greater degree of non-compliance, with the remaining guideline "Distinguish required and optional data entry fields" often not being applicable because of a lack of data entry forms on a site. In these cases it was scored as partially complied so as to not skew data.

The Correlation Between Usability and Web Site Popularity

The main objective of this study was to measure the strength and significance of the correlation between compliance with usability guidelines and Web site popularity. Using Spearman's Rank method, the correlations between the Usability Score of each reviewed Web site and each of

the five ranks that were included in the data collection process were measured. Appendix Table 2 shows the scores of the reviewed sites prior to testing for correlation.

As previously established, the most important measure of popularity for the purpose of this study is the number of inbound links to the Web site in Google. The other popularity measurements have been included to further support the findings. The first analysis measured the correlations across all 38 sites included in the study. The results are shown in Appendix Table 3.

When measuring the correlation between compliance to usability guidelines with the five ranking systems across all 48 Web sites a significant correlation is visible in all cases except for that of Del.icio.us online bookmarking. The correlation between Usability Score and links in Google is of medium strength (.594) and is significant at the 0.01 level (2-tailed). This is the strongest correlation between Usability Score and the ranking systems.

High correlations are also visible between the number of Google links and Google PageRank and Yahoo link-count (.797 and .792 respectively). Both of these correlations are also significant. Medium correlations between Google link-count and the Alexa rating and number of Del.icio.us bookmarks are also visible. Because the Alexa system ranks Web sites from 1, the most visited Web site on the net, to an infinite number, the least-visited, the correlation between Alexa and other scores appears negative.

To counter the affect that the subject matter of a Web sites being reviewed may have on its popularity, or potential reach, the sites were grouped by subject and the same tests were conducted on two of the groups. The larger of these groups was Web sites that contain educational resources for a diverse range of subjects. Appendix Table 4 presents the findings of tests on this group of 10 sites.

Again, the most significant and strongest correlation between Usability Score and a ranking system is that of Google link-count. Having narrowed down the sample to a group of sites

with similar content, a very strong correlation becomes visible. Furthermore, despite the small sample size, this correlation coefficient of .855 is significant at the 0.01 level (2-tailed). Also visible in this smaller sample is a strong correlation (.830) between the Usability Score and Google PageRank. Other than a slight decrease in the correlation between Google link-count and Google PageRank and Yahoo link-count, the other correlations visible in this test are similar or even stronger than with the larger sample.

The same patterns emerge in the 3rd test, that of Web sites that provide educational resources for business courses. With a sample of just six sites, a significant correlation of .941 is found between compliance to usability guidelines and the number of links pointing to the site in Google. In this sample, strong correlations between the pairs Usability Score - Google PageRank and Google link-count - Google PageRank are also shown (Appendix Table 5). None of the six sites had any bookmarks on Del.icio.us, resulting in no correlation data being calculable.

All of the correlations visible in these three tests provide support for the hypothesis of this study, that compliance with usability guidelines has a positive effect on Web site popularity.

CONCLUSION

Observed Compliance with Usability Guidelines

Overall the 28 usability guidelines selected for the study are complied with more frequently than not. For the most part, it seems that usability is either being considered, or it is naturally occurring, in Web site development. However, simple guidelines which really seem fundamentally obvious in design are visibly ignored or forgotten in many cases. How can it be that over half of the sites reviewed had cluttered displays on some or all of their pages? Or that only 60.5% of sites display the majority of their content in high-contrast color combinations so that it can actually be read?

Some of the guidelines show such high levels of compliance that it can be suggested they are now common-sense Web site design rules. Only one Web site displayed a horizontal scrollbar, which indicates that the problem of horizontal scrolling has been almost eliminated from the Web. It appears that Web site developers have also achieved a greater command of English, in that prose text is nearly always formatted correctly in mixed-case.

Page download time remains as much of an issue today as it did 10 years ago. Less than 14% of the sites reviewed downloaded in less than 10 seconds of a 56k modem, leaving the remaining sites at risk of losing visitors. It is possible the developers of the slow sites have calculated that with the uptake of broadband technologies the risk of upsetting visitors is too minimal to be of concern. But this leaves all visitors on modem connections in an unfortunate predicament. Should these users be dictated to and forced into more expensive technology by Web developers?

Increased Popularity Through Usability

All five of the ranking methods that data were collected for show positive correlations with compliance to usability guidelines. The main indicator of popularity in this study, the number of inbound links found in Google, consistently shows a strong to very strong correlation across all three tests. Breaking the Web sites down into groups based on the content subject achieved a stronger result. The lack of a more complete sample for this study creates the need for further studies of this nature to be completed. However, we are confident that the same results obtained now would be reflected in the results from these tests over a sample of any size.

We believe that the principle hypothesis of this study has been proven, that is; increased compliance with usability guidelines does have a correlation with increased popularity of a Web site. Moreover, this correlation is shown to be both very strong and significant. Causality cannot be implied by a correlation alone,

however the indication is indeed that Web sites that comply with usability guidelines will also inherently be more accepted by users, and thus boast higher popularity.

More usable Web sites not only acquire a greater number of links from other Web sites, but they also achieve both a higher Google PageRank and are more popular according to the Alexa rating system. Perhaps, then, more focus might be given to usability as a means to achieving success in Web sites in the future. Further research of this nature is called for, to enhance and support the results obtained so far. If the results of this study can be confirmed, perhaps Web site usability consulting services will become as popular as search engine optimization has been in the last few years.

Ranking Systems

Although the review of literature suggests that the Alexa ranking system is biased and not a reliable indicator of Web site popularity, it is apparent that it is at least consistent with the other rankings used. In all cases, the Alexa rank showed the appropriate negative correlation with the other ranking systems and the Usability Score.

The lack of Del.icio.us bookmarks came as something of a surprise. A system suggested to have been an extremely popular method of sharing Web sites turned out to be disappointing. Only 5 of the sites reviewed had any bookmarks in the Del.icio.us system. The majority of these sites did also exhibit high counts of links on Google and Yahoo, but it still brings a few questions to mind. Is Del.icio.us as popular as rumors suggest? If it is, then what sort of elitism does a site need to achieve before it appears in Del.icio.us? Does the average internet user want to bookmark their favorite sites online, or do they even know they can? And finally, do users of Del.icio.us fall into any particular age or demographic that results in favoritism towards certain types of site? In our opinion, comparing the number of bookmarks in Del.icio.us is not a suitable way of determining popularity.

Google PageRank cannot be considered to be a measure of popularity, but it is intended to indicate quality in a site. Compliance with usability guidelines correlates to increased popularity but also that increased usability results in popularity among higher quality referrers.

The findings of the study show a strong and significant correlation between the quantities of links found on Google and Yahoo, suggesting that Yahoo link-count may also be a suitable indicator of popularity. Yahoo, however, whilst always showing a correlation between Usability Score and link-count, did not consistently show a significant correlation. The number of links on Yahoo was normally many times that of Google and often featured links from partner sites in a related network (link farms) or from within the site being explored. Where Google has clearly spent effort developing a system for scoring the quality of links (PageRank) and minimizing the quantity of spam links, Yahoo has failed. This has resulted in the quality of Yahoo link-count as an indicator of popularity being diminished.

Limitations of this Study

The principle limitation of this study is the quantity and quality of the sample. Future studies need to be conducted with a larger sample which has been more carefully sourced. Following the processing of the original sample, only 38 Web sites remained from an original count of over 200. This number of Web sites, while sufficient enough to establish results, did not constitute a coherent sample.

Furthermore, nine of the sites reviewed were location-centric, six to the UK and three to New Zealand. The New Zealand based sites, in particular, were outliers; often with a high Usability Score but a low popularity due to the limited size of their target audience. In future studies this issue should also be avoided through careful sampling.

Finally, the snapshot nature of this research is a possible limitation. What is considered vital to Web site success one year can change in the next. The correlation between usability

and popularity can therefore also change at a rapid pace. Research of this nature should be made on a regular basis, perhaps annually, to determine if these results are indicative of the Internet industry through time, or just a one-off occurrence.

Discussion

This study shows that certain usability guidelines are complied with frequently, while others are almost completely ignored. What does this mean for usability practitioners? Is it possible that some of the guidelines published are too hard to implement, or that developers feel too limited by them? Should guidelines that have become common-sense be published at all?

Certainly the suggestion that increased compliance with usability guidelines has a strong relationship with increased Web site popularity is of importance to the Internet industry. Will further findings of this type result in a higher uptake of usability studies during Web site development? The implications of this finding for organizations developing Web sites is if they want to attract more visitors, then some effort could be well applied to usability awareness. For usability consulting companies, this research calls for more studies of a similar nature to help strengthen the industry and to build greater awareness of the importance of usability and possible benefits.

ACKNOWLEDGMENT

This article is based on the first author's unpublished case study thesis in Information Science at the University of Otago. We would like to thank all our colleagues for their advice and, in particular, Colin Aldridge for his methodological guidance. Thanks to Keryn Pratt for providing the necessary e-learning resources and her help.

REFERENCES

- W3C. (2001). *WCAG WG minutes 29 March 2001*. Retrieved July 20, 2007, from <http://www.w3.org/WAI/GL/2001/03/29-minutes.html>
- Adamic, L. A., & Huberman, B. A. (2001). The Web's hidden order. *Communications of the ACM*, 44(9), 55–60. doi:10.1145/383694.383707
- Alexa. (2007). *Alexa the Web Information Company*. Retrieved May 7, 2007, from <http://www.alexa.com>
- Belanger, F., Fan, W., Schaupp, C. L., Krishen, A., Everhart, J., & Poteet, D. (2006). Web site success metrics: Addressing the duality of goals. *Communications of the ACM*, 49(12), 114–116. doi:10.1145/1183236.1183256
- Brajnik, G. (2000). *Automatic Web usability evaluation: What needs to be done?* Retrieved May 5, 2007, 2007, from <http://www.dimi.uniud.it/~giorgio/papers/hfweb00.html>
- Brin, S., & Page, L. (1998, April 14-18). *The anatomy of a large-scale hypertextual Web search engine*. Paper presented at the International Conference on World Wide Web, Brisbane, Australia.
- Cho, J., & Roy, S. (2004, May 17-20). Impact of search engines on page popularity. In *Proceedings of the 13th International Conference on World Wide Web*, New York (pp. 20-29). ACM Publishing.
- De Angeli, A., Sutcliffe, A., & Hartmann, J. (2006, June 26-28). Interaction, usability and aesthetics: What influences users' preferences? In *Proceedings of the 6th ACM Conference on Designing Interactive Systems*, University Park, PA (pp. 271-280). ACM Publishing.
- Del.icio.us. (2007). *del.icio.us*. Retrieved May 7, 2007, from <http://del.icio.us/>
- Gladwell, M. (2000). *The tipping point: How little things can make a big difference*. New York: Time Warner Book Group.
- Google. (2007). *Advanced Google search operators*. Retrieved May 7, 2007, from <http://www.google.com/intl/en/help/operators.html>
- HHS. (2006). *Research-based Web design & usability guidelines*. Retrieved March 13, 2007, from <http://www.usability.gov/pdfs/guidelines.html>
- Instone, K. (1999). *How to test usability*. Retrieved from <http://usableweb.com/instone/howtotest>

- ISO. (1998). *International Standard ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) - part 11: Guidance on usability*. Retrieved March 13, 2007, from <http://www.idemployee.id.tue.nl/g.w.m.rauterberg/lectures.html>
- Ivory, M. Y., & Megraw, R. (2005). Evolution of Web site design patterns. *ACM Transactions on Information Systems*, 23(4), 463-497. doi:10.1145/1095872.1095876
- Kavassalis, P., Lelis, S., Rafea, M., & Haridi, S. (2004). What makes a Web site popular? *Communications of the ACM*, 47(2), 50-55. doi:10.1145/966389.966415
- Keevil, B. (1998, September 24-26). Measuring the usability index of your Web site. In *Proceedings of the 16th Annual International Conference on Computer documentation*, Quebec City, Quebec, Canada (pp. 271-277). ACM Publishing.
- Koyani, S., & Allison, S. (2003, April 5-10). Use of research-based guidelines in the development of Web sites. In *Extended Abstracts on Human Factors in Computing Systems: Proceedings of CHI '03*, Ft. Lauderdale, FL (pp. 696-697). ACM Publishing.
- Lederer, A. L., Maupin, D. J., Sena, M. P., & Zhuang, Y. (1998, March 26-28). The role of ease of use, usefulness and attitude in the prediction of World Wide Web usage. In *Proceedings of the 1998 ACM SIGCPR Conference on Computer Personnel Research*, Boston (pp. 195-204). ACM Publishing.
- McDonald, J. (2006). *Spearman's rank correlation*. Retrieved September 9, 2007, from <http://www.udel.edu/~mcdonald/statspearman.html>
- Nielsen, J. (1993). *Usability engineering*. San Diego, CA: Academic Press.
- Nielsen, J. (1999). User interface directions for the Web. *Communications of the ACM*, 42(1), 65-72. doi:10.1145/291469.291470
- Nielsen, J., & Mack, R. (1994). *Usability inspection methods*. New York: Wiley.
- Ohnemus, K. R. (1997, October 19-22). Web style guides: Who, what, where. In *Proceedings of the 15th Annual International Conference on Computer Documentation*, Salt Lake City, UT (pp. 189-197). ACM Publishing.
- Oxford. (n.d.). *Compact Oxford English dictionary (ask Oxford)*. Retrieved July 08, 2007, from http://www.askoxford.com/concise_oed/popular?view=uk
- Petrie, H., & Kheir, O. (2007, April 28-May 3). The relationship between accessibility and usability of Web sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, San Jose, CA (pp. 397-406). ACM Publishing.
- Schonberg, E., Cofino, T., Hoch, R., Podlaseck, M., & Spraragen, S. L. (2000). Measuring success. *Communications of the ACM*, 43(8), 53-57. doi:10.1145/345124.345142
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72-101. doi:10.2307/1412159
- Spiliopoulou, M. (2000). Web usage mining for Web site evaluation. *Communications of the ACM*, 43(8), 127-134. doi:10.1145/345124.345167
- Sullivan, T., & Matson, R. (2000, November 16-17). Barriers to use: Usability and content accessibility on the Web's most popular sites. In *Proceedings on the 2000 Conference on Universal Usability*, Arlington, VA (pp. 139-144). ACM Publishing.
- The Nielsen Company. (2007). *Nielsen netratings*. Retrieved July 9, 2007, from <http://www.nielsen-netratings.com>
- Theofanos, M. F., & Redish, J. (2003). Bridging the gap: Between accessibility and usability. [ACM]. *Interaction*, 10(6), 36-51. doi:10.1145/947226.947227
- Usablenet. (n.d.). *LIFT online*. Retrieved April 12, 2007, from http://www.usablenet.com/products_services/lift_online/lift_online.html
- Wattenberg, T. (2004). Beyond standards: Reaching usability goals through user participation. *SIGACCESS Accessibility and Computing*, 79, 10-20. doi:10.1145/1040053.1040055
- Weinschenk, S. (2007). *Trends in user-centred design* (White Paper). Fairfield, IA: Human Factors International.
- Yahoo. (n.d.). *Yahoo site explorer*. Retrieved May 7, 2007, from <http://siteexplorer.search.yahoo.com>

APPENDIX

Table 1. Compliance to usability guidelines in 38 Web sites

#	Guideline	Complied		Partially Complied		Not Complied	
		#	%	#	%	#	%
1	Do not display unsolicited windows or graphics	33	86.8	4	10.5	1	2.6
2	Show all major options on the homepage	34	89.5	2	5.3	2	5.3
3	Avoid cluttered displays	16	42.1	8	21.1	14	36.8
4	Place important items consistently?	21	55.3	8	21.1	9	23.7
5	Eliminate horizontal scrolling	36	94.7	1	2.6	1	2.6
6	Use meaningful link labels	22	57.9	14	36.8	2	5.3
7	Distinguish required and optional data entry fields	10	26.3	16	42.1	12	31.6
8	Label pushbuttons clearly	18	47.4	17	44.7	3	7.9
9	Organise information clearly	26	68.4	9	23.7	3	7.9
10	Facilitate scanning	19	50.0	7	18.4	12	31.6
11	Ensure that images do not slow downloads	25	65.8	3	7.9	10	26.3
12	Include logos	31	81.6	4	10.5	3	7.9
13	Minimize page download times	5	13.2	0	0.0	33	86.8
14	Provide text equivalents for non-text elements	14	36.8	5	13.2	19	50.0
15	Provide a search option on every page	17	44.7	1	2.6	20	52.6
16	Communicate the websites value and purpose	20	52.6	12	31.6	6	15.8
17	Limit homepage length	19	50.0	2	5.3	17	44.7
18	Use bold text sparingly	27	71.1	3	7.9	8	21.1
19	Provide feedback on user's location	14	36.8	4	10.5	20	52.6
20	Provide descriptive page titles	16	42.1	9	23.7	13	34.2
21	Use descriptive headings liberally	19	50.0	13	34.2	6	15.8
22	Link to related content	32	84.2	3	7.9	3	7.9
23	Use text for links	26	68.4	11	28.9	1	2.6
24	Use black text on plain, high-contrast backgrounds	23	60.5	8	21.1	7	18.4
25	Use mixed-case for prose text	36	94.7	2	5.3	0	0.0
26	Ensure visual consistency	27	71.1	6	15.8	5	13.2
27	Format lists to ease scanning	26	68.4	4	10.5	8	21.1
28	Label data entry fields consistently	24	63.2	13	34.2	1	2.6
	Mean	23	59.8	7	17.8	9	22.5
	Median	23	59.8	6	15.8	7	18.4
	Maximum	36	94.7	17	44.7	33	86.8
	Minimum	5	13.2	0	0.0	0	0.0

Table 2. Review scores (by subject then usability score)

Site	Alexa Rating	PageRank	GoogleLinks	YahooLinks	DeLicio.us	Usability Score	Loc	Subject
022	2478826	5	15	659	0	107.5		Astronomy
036	338704	5	159	3768	6	96	NZ	Biology
028	235941	6	141	5833	0	92.5		Biology
027	70158	6	290	31151	644	66.5		Biology
038	85387	7	732	7355	0	116	UK	Business
034	0	3	5	167	0	87	NZ	Business
030	1126005	5	87	2136	0	83.5		Business
035	502470	4	56	2655	0	79		Business
033	168971	4	61	5834	0	68		Business
032	0	4	3	297	0	65.5		Business
029	5899631	0	0	147	0	63		Business
052	1325979	4	4	743	0	71		Classics
023	1360	8	38100	737123	12989	107		Diverse
014	306470	7	783	28677	0	105.5		Diverse
007	2360	7	614	1459	0	101		Diverse
024	42	7	1140	13340	0	98	UK	Diverse
019	23057	6	1720	59284	5	87.5		Diverse
001	117478	1	413	1942	0	83		Diverse
004	2184730	5	81	1708	0	81		Diverse
042	142425	5	136	9813	0	75.5	UK	Diverse
005	112354	5	104	1580	0	59.5	UK	Diverse
050	1026411	5	20	7180	0	58.5	UK	Diverse
017	229847	6	619	25517	0	99.5		English
031	877519	6	147	1033	0	66		Ethics
046	8601318	4	94	6918	0	60.5		Ethics
037	815276	6	320	322	0	88.5		History
053	175772	6	314	14499	0	75.5		History
039	605	8	5700	242450	31000	113.5		Internet
040	14542	7	5790	1071387	1629	94.5		Internet
048	399853	6	488	13606	0	75		PE
047	1455426	4	15	934	0	46.5		PE
049	5542429	5	43	1494	0	39		PE
043	91802	6	144	18972	0	101.5		Physics
041	397750	5	6	142	0	92.5	NZ	Physics
044	437826	7	276	6221	0	76		Physics
008	2230007	7	405	15821	0	65		Science
021	27148	6	992	50207	0	104		Technology
016	455728	5	679	8781	0	102	UK	Technology

Table 3. Spearman's rho correlations (all sites)

		Alexa	PageRank	GoogleLinks	YahooLinks	Delicious	Usability Score
Alexa	Correlation Coefficient	1.000	-.379(*)	-.490(**)	-.392(*)	-.418(**)	-.456(**)
	Sig. (2-tailed)	.	.019	.002	.015	.009	.004
	N	38	38	38	38	38	38
PageRank	Correlation Coefficient		1.000	.797(**)	.657(**)	.391(*)	.527(**)
	Sig. (2-tailed)		.	.000	.000	.015	.001
	N		38	38	38	38	38
GoogleLinks	Correlation Coefficient			1.000	.792(**)	.478(**)	.594(**)
	Sig. (2-tailed)			.	.000	.002	.000
	N			38	38	38	38
YahooLinks	Correlation Coefficient				1.000	.530(*)	.366(*)
	Sig. (2-tailed)				.	.001	.024
	N				38	38	38
Delicious	Correlation Coefficient					1.000	.282
	Sig. (2-tailed)					.	.086
	N					38	38
Usability Score	Correlation Coefficient						1.000
	Sig. (2-tailed)						.
	N						38

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

Table 4. Spearman's rho correlations (diverse subject sites)

		Alexa	PageRank	GoogleLinks	YahooLinks	Delicious	Usability Score
Alexa	Correlation Coefficient	1.000	-.608	-.758(*)	-.261	-.450	-.552
	Sig. (2-tailed)	.	.062	.011	.467	.192	.098
	N	10	10	10	10	10	10
PageRank	Correlation Coefficient		1.000	.754(*)	.500	.488	.830(**)
	Sig. (2-tailed)		.	.012	.141	.152	.003
	N		10	10	10	10	10
GoogleLinks	Correlation Coefficient			1.000	.697(*)	.701(*)	.855(**)
	Sig. (2-tailed)			.	.025	.024	.002
	N			10	10	10	10
YahooLinks	Correlation Coefficient				1.000	.701(*)	.479
	Sig. (2-tailed)				.	.024	.162
	N				10	10	10
Delicious	Correlation Coefficient					1.000	.467
	Sig. (2-tailed)					.	.173
	N					10	10
Usability Score	Correlation Coefficient						1.000
	Sig. (2-tailed)						.
	N						10

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

Table 5. Spearman's rho correlations (business subject sites)

		Alexa	PageRank	GoogleLinks	YahooLinks	Delicious	Usability Score
Alexa	Correlation Coefficient	1.000	-.395	-.257	-.429	.	-.200
	Sig. (2-tailed)	.	.439	.623	.397	.	.704
	N	6	6	6	6	6	6
PageRank	Correlation Coefficient		1.000	.941(**)	.698	.	.941(**)
	Sig. (2-tailed)		.	.005	.123	.	.005
	N		6	6	6	6	6
GoogleLinks	Correlation Coefficient			1.000	.829(*)	.	.943(**)
	Sig. (2-tailed)			.	.042	.	.005
	N			6	6	6	6
YahooLinks	Correlation Coefficient				1.000	.	.771
	Sig. (2-tailed)				.	.	.072
	N				6	6	6
Delicious	Correlation Coefficient					1.000	.
	Sig. (2-tailed)					.	.
	N					6	6
Usability Score	Correlation Coefficient						1.000
	Sig. (2-tailed)						.
	N						6

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

Greg Scowen received his bachelor's degree in information science from the University of Otago in 2006 and completed post-graduate research studies at the University of Otago in 2007. He is currently working as a software engineer/Web developer at the Swiss Federal Institute of Technology (ETH) in Zürich, developing advanced Web applications for the library. His research interest areas include usability, Web development, and e-learning.

Holger Regenbrecht is a senior lecturer at the department of information science at Otago University. His research interests include human-computer interaction (hci), (collaborative) augmented reality, 3D teleconferencing, psychological aspects of mixed reality, and three-dimensional user interfaces (3DUJ). He is a member of IEEE, ACM, and igroup.org and serves as a reviewer and auditor for several conferences and institutions.