An Inclusive Guide
To Assessing Web Site Effectiveness
Janna B. Arney, (Email: jbarney@utb.edu), University of Texas at Brownsville
Paul J. Lazarony, California State University, Northridge

ABSTRACT
This study presents the results of a literature review of the well-established literatures on readability, usability, and Web design. The purpose of this study is to provide a comprehensive review of the literature on readability, usability, and Web design and to propose a three-step evaluation process to assess Web sites in each of these three areas. We recommend that this process be used to evaluate existing Web documents and those still in the design phase. This evaluation process can be utilized as an educational tool in a classroom setting. Using this process in a classroom provides students with much needed training in a real-world approach to Web site testing without monopolizing the course content. The research is presented in four sections: (i) Readability, (ii) Usability, (iii) Web Design Guidelines, and (iv) Three-Step Evaluation Process.

INTRODUCTION
Despite its highly touted capabilities, the World Wide Web continues to generate substantial frustration among users (Palmer, 2002). Web designers, both professionals and students, have access to an abundance of design guidelines. Effective Web designers follow strict design guidelines to ensure that their Web sites meet the needs of their users. Although following these guidelines is of the utmost importance in designing a Web site, of equal importance is assessing whether these guidelines have been successfully implemented from a user’s perspective. As educators, we have an obligation to instruct students in proper Web design but also to equip them with evaluation tools to assess the effectiveness of Web sites, their own and those of others. To this end, the purpose of this study is to provide a review of the literature on readability, usability, and Web design and to propose a three-step evaluation process to assess Web sites in each of these three areas. This evaluation process can be successfully utilized as an educational tool in a classroom setting. The next three sections define readability, usability, and Web design guidelines and describe methods for assessing each.

READABILITY
Readability means writing so that audience members can read and understand the message. Readability and understandability are not always the same. According to Leong, Ewing, and Pitt (2002), readability is text-centered while understandability is reader-centered. Readability has to do with the way the text is written and presented to the reader whereas understandability includes reader characteristics such as reader competence and reader motivation. Although understandability of a reader cannot be easily predicted by an author, readability can and should be addressed because of its importance in measuring the level of difficulty of textual information.

“With the multiplicity and diverse variety of web pages containing information, comprehension becomes a critical issue, and a major facet of comprehension is readability of content” (Leong, et al., 2002). A particular challenge in designing Web materials is to identify the composition of the audience. Although it is impossible to assess every reader’s competence and motivation, it is possible to measure the difficulty of the text in an effort to reach as many audience members as possible.
According to Klare (1974), an author has several ways in which to approach the readability problem: guess at the reading level of the audience; develop and administer a comprehension test covering the material to a group of readers; or use "readability" formulas. This last option is the most practical and efficient because it may easily be applied to the large amounts of information published by individuals and companies. Two readability formulas are discussed in the next sections: Flesch Reading Ease and Flesch/Kincaid Grade Level. Both of these measures can be calculated using Microsoft Word.

**Flesch Reading Ease**

The Flesch Reading Ease formula has become one of the most widely used formulas in the history of readability measurement (Klare 1974). The result of applying the Reading Ease formula to a sample of text is a rating ranging from 0 to 100, with higher ratings indicating material that is easier to read, as shown Table 1.

**Flesch/Kincaid Grade Level**

Kincaid, Fishburne, Rogers, & Chissom (1975) followed the Flesch Reading Ease formula with the Flesch-Kincaid Grade Level formula. Like the Flesch Reading Ease formula, the Flesch-Kincaid Grade Level formula computes readability based on the average number of syllables per word and the average number of words per sentence. The computation results in a score that can be interpreted as a grade level. The grade levels computed using Flesch-Kincaid are good predictors of reading difficulty (Clariana, 1993). Table 2 has the results of a sample calculation where the level of 6.0 indicates that the passage of text was written at the 6th grade level.

Both of these readability formulas can be applied to the textual information on Web pages which can provide valuable information on Web design. However, readability alone cannot provide a complete picture as to the effectiveness of a Web page. The next section introduces usability. Both readability and usability principles should be followed to judge effective Web sites. Web designers, unlike writers, must be concerned with the users’ background or previous experience with the Web as a communication medium. “The architecture, interface, and interaction need to be at the same level as the users” (Zibell, 2000). Like writers, Web designers do not have the luxury of knowing the level of knowledge of every user. This added challenge of Web-based documents is usability. Usability and readability become synonymous when designing a readable yet usable Web site. “If a user can’t use or read the information, or even find it, then it might as well not be” (Zibell, 2000).

**USABILITY**

“Usability is a quality attribute that assesses how easy user interfaces are to use. The word usability also refers to methods for improving ease-of-use during the design process” (Nielsen, 2003). Usability primarily involves “what elements appear onscreen and how efficient, intelligible, and intuitive they are” (Palmer, 2002). Whether a Web site is successful depends on how usable it is perceived to be by users.

To determine whether your Web site is likely to be perceived as usable, you must first evaluate it. Usability can be evaluated in a variety of ways, from sophisticated and expensive to simple and informal. “The method you choose and the depth at which you conduct usability testing should correlate to the potential risk, such as lost revenue, associated with poor usability” (Guenther, 2003). In other words, if the potential of revenue loss is great so should be the investment in usability testing.

“The most expensive, but arguably the most thorough, option for usability testing is to outsource the testing with an outside vendor” (Guenther, 2003). According to such a vendor, Software Usability Research Laboratory (2003), usability testing is an empirical method of measuring a Web site’s ease-of-use by bringing representative users into a fully-equipped usability laboratory. The users are asked to complete a series of tasks with the Web site. Data is collected by observing the users, by soliciting feedback on user satisfaction, and by collecting performance data. The goal is to assess the Web site’s ease of use, efficiency, usefulness, and appeal to its users.
For those electing to conduct usability testing independent of vendors but with the desire to accomplish the same goals, two usability evaluation methods are available: heuristic evaluation and end-user evaluation. The first method, heuristic evaluation, is a usability engineering method that is cheap, fast, and easy to use (Nielsen, 1994). It is useful in identifying usability problems during the design phase of user interfaces, thereby allowing for the correction of potential problems prior to the implementation of a system. The second method, end user evaluation, determines how easily typical users can accomplish tasks that are critical to the success of the Web site (Lisney & Schang, 2001). Regardless of the method selected, “engaging end users throughout the process of developing a Web site is critical to the success of your project. Doing so has significant advantages, namely the assurance that your project is on track during all phases of development, not just at the end when it comes time to formally test the product” (Guenther, 2003).

Heuristic Evaluation

Nielsen’s heuristic evaluation requires a team of three to five evaluators to examine the interface and then to judge its compliance with recognized usability principles or the “heuristics.” Nielsen recommends a team of three to five since significant gain has not been reported with larger groups. In fact, according to King (2003), testing five people three times will provide more and better information than testing 15 people once. Once all the evaluations have been completed, the results are aggregated. According to Nielsen (1994), heuristic evaluations can detect between 40 and 60 percent of the usability problems that an empirical user test would find, with the types of problems being roughly comparable.

User Testing

According to a more recent study by Nielsen (2003), user testing is recommended above all other methods for studying usability; he describes it as having “three components:

1. Get hold of some representative users, such as customers for an e-commerce site or employees for an intranet.
2. Ask the users to perform representative tasks with the design.
3. Observe what the users do, where they succeed, and where they have difficulties with the user interface. Shut up and let the users do the talking.”

How do you determine who your representative users are? Some examples of users selected for other studies are Boiarsky (2002) where 45 English students were given questionnaires to evaluate 12 B2C Web sites selected by the author; Website Usability Testing Center at The University of Wisconsin-Stout (2001) where high school students were selected as subjects to evaluate the university’s Web site; Linn (2001) where eight heuristic groups evaluated six university Web sites; King (2003) where five volunteers (library patrons) were selected to evaluated the Kansas City Public Library Web site; and Lisney & Schang (2001) where nine participants were selected to evaluate the Any Baby Can Web site followed by two Web site usability experts who conducted a heuristic evaluation.

To assess usability from a user’s perspective, Palmer (2002) designed an instrument to measure Web site usability, design, and performance. “The result of his research effort is a set of parsimonious, understandable, and reliable metrics that have been tested in multiple studies, with different user groups, across a multiyear period” (Palmer, 2002). Specifically, Palmer (2002) hypothesized that five factors would be associated with greater perceived success by site users: lower download delay, more navigable Web sites, higher interactivity, more responsive Web sites, and higher quality content. To test these hypotheses, three studies were completed during a three-year period, from 1997 to 2000. Each subsequent study used the metrics identified and refined in the prior study but then added an additional data gathering method (Palmer, 2002). The resulting instrument was tested for reliability, convergent and construct validity, nomological validity, external validity, and generalizability (Palmer, 2002). The result is a user-friendly evaluation instrument that can be used in a variety of settings and can be customized based on individual needs.
Another instrument designed to evaluate the usability from the user perspective was developed by the Website Usability Testing Center at the University of Wisconsin-Stout (UW-Stout). This instrument was developed specifically to evaluate usability from the perspective of a potential student. The study selected potential high school students as representative users who are unfamiliar with the university Web site. “As infrequent users, the site must be usable and impressive to portray a positive image of the University” (Website Usability Testing Center, 2001). The research staff from the UW-Stout Website Usability Testing Center created a usability checklist of generally accepted principles. Evaluators were asked to evaluate the university Web site checking that these principles had been followed. For the user testing, a questionnaire was developed with the following sections: (1) User Background and Experience, (2) West Site Information Retrieval Tasks, and (3) Web Site Evaluation. The Web Site Evaluation section asked students to rate the Web site using a Likert Scale on attributes such as load time, system of navigation, general organization, visuals, and overall site rating. Students were also asked how they would compare this university’s site to that of other universities’ Web sites.

Based on the three studies discussed, Nielsen, Palmer, and the Website Usability Testing Center, an evaluation instrument was developed for this study (See Appendix A). Once a Web site has been evaluated for readability and usability, it must also be judged against a checklist of generally accepted design guidelines.

WEB DESIGN GUIDELINES

According to Palmer (2002), “businesses whose homepages address usability and incorporate other essential design criteria report higher traffic, more repeat visitors, and greater customer satisfaction.” Usability and design become one in the same when designing an effective Web site. Although much has been written about Web design, it is often to the exclusion of usability, which is significantly associated with Web design. In this section, the most frequently described design guidelines are presented.

Palmer presents four basic design elements that must always be considered (2002).

1. Navigation – Design should aid users in finding information quickly and efficiently.
2. Response time – Design should provide users with fast loading pages and readily available search results.
3. Content – Design should constantly update company information, organize content within a single click where possible, offer a number of printing options, and order the most relevant material at the top of a page and at the top of the hierarchical structure.
4. Interactivity and responsiveness – Design should give users the opportunity to customize their interactions and should provide users with appropriate feedback.

Often we can learn from others’ mistakes as well as from their successes. This is likely why Nielsen (2003) published the Ten Most Violated Homepage Design Guidelines. According to Nielsen (2003), these very important guidelines are not followed by two-thirds of corporations who publish Web sites. The guidelines listed below are sorted in the order of violation, with the most commonly violated guideline listed first (Nielsen, 2003).

1. Emphasize what your site offers that’s of value to users and how your services differ from those of key competitors.
2. Use a liquid layout that lets users adjust the homepage size.
3. Use color to distinguish visited and unvisited links.
4. Use graphics to show real content, not just to decorate your homepage.
5. Include a tag line that explicitly summarized what the site or company does.
6. Make it easy to access anything recently featured on your homepage.
7. Include a short site description in the window title.
8. Don’t use a heading to label the search area; instead use a “Search” button to the right of the box.
9. With stock quotes, give the percentage of change, not just the points gained or lost.
10. Don’t include an active link to the homepage on the homepage.
In addition to correcting the ten most violated homepage design guidelines, Nielsen reminds us to beware of the Top Ten Web-Design Mistakes of 2002 (Nielsen, 2002). These mistakes have to do with general Web page design not just homepage design. The top ten mistakes are no prices published (ecommerce sites); inflexible search engines; horizontal scrolling – Sites must not require users to scroll left to right; fixed font size rather than an absolute number of pixels; blocks of text; JavaScript in links – links should be a simple hypertext reference; infrequently asked questions in FAQ; collecting email addresses without a privacy policy; URL > 75 characters; mailto links in unexpected locations.

According to the Website Usability Testing Center (2001), a checklist of design guidelines must first be used to evaluate a Web site. The design guidelines determined most important by the research staff were load time, navigation, structure/layout, content, and visuals. These agree with the research of Palmer and Nielsen. In fact, the body of literature on Web design guidelines is quite predictable because of its consistency. However, in spite of the agreement on what constitutes good Web design, an infinite number of Web sites are published that violate these basic design principles. To test for these principles as well as the readability and usability principles discussed earlier, the next section outlines a three-step evaluation process.

THREE-STEP EVALUATION PROCESS

Student Web designers should consider all three categories of evaluation discussed: (1) readability, (2) usability, and (3) design guidelines. In the next sections, we recommend a three-step evaluation process.

Step 1: Evaluate Readability

Readability is the first and easiest step in the evaluation process. To determine the readability of a passage of text, you must complete a series of steps.

1. Copy the textual information to be included on a Web page and paste it into a Microsoft Word document.
2. Verify that the readability statistics option is selected in your software. Select Tools then Options then the Spelling and Grammar tab. In the Grammar section, ensure that the Show Readability Statistics option is checked. Select OK.
3. Run the Grammar Checker. Select Tools then Spelling and Grammar. Once the Grammar Checker has completed checking the spelling and grammar of the text, you can elect to view the Readability Statistics. Sample readability results can be found in Table 2.

The readability of a passage of text gives you only a rough idea of how well the audience will understand your message. Readability formulas measure the length of words and sentences. Bovée, Thill, & Schatzman (2003) suggest writing at a 12th to 14th-grade level for technical documents; for general business messages, however, your writing should be aimed at the 8th- to 11th-grade level. It is important to note that the readability statistics generated by Microsoft Word do not score any passage of text higher than the 12th-grade level.

If the readability statistics indicate a grade level of 10th or higher, begin revising by checking sentence length and the difficulty of your words and then simplifying where possible. “Readability research tells us that the more words and the more relationships there are in a sentence, the greater is the possibility for misunderstanding” (Lesikar, Pettit, & Flatley, 1999). Shorter sentences and simpler words are better understood by the average reader. Therefore, revise and recalculate until you believe that your writing is appropriate for your intended audience. “The value of calculating a readability measure lies in the valuable feedback you gain” (Lehman & Dufrene, 2002). Although readability formulas can provide valuable feedback, they cannot provide a complete picture as to the effectiveness of Web documents. In spite of their usefulness, “they ignore some important variables that also contribute to reading ease, such as sentence structure, the organization of ideas, and the appearance of the message on the page” (Bovée et al., 2003). To address these variables, good writers follow the standard business communication principles for effective writing. In addition, when producing Web materials, good writers/designers also consider usability principles and Web design guidelines.
Step 2: Evaluate Usability

The second step in the evaluation process is to test the usability of your Web site. Usability is one of the most time-consuming but arguably the most important in the evaluation process. According to Guenther (2003), to conduct your own usability testing, you must identify and recruit a representative user group; design the test; include tasks to be performed by users; determine criteria to be measured along with the instrument to collect user data; develop the necessary materials; prepare the test environment; conduct the test; document the findings; and analyze the user data to determine what changes or enhancements need to be made to the Web site. Based on the research of Guenther (2003), Nielsen (2003), and King (2003), we recommend completing the following steps to determine the usability of your Web site. These can be applied to a Web site under construction or to the redesign of an existing Web site. These can be successfully used in the classroom by allowing teams to evaluate each other’s team projects.

1. Identify the purpose of your Web site and then of each page within the site. As with business writing, defining your general purpose must be followed by defining your specific purpose. For example, you are charged with redesigning your departmental Web site. First, define your general purpose, “to make the site more user-friendly.” Then define several specific purposes for each Web page, “to make the degree plan information easier to find and easier to print.”

2. Based on the purpose, derive a list of task questions. The questions should be specific, not open-ended. Following the same department Web site redesign example, a sample question would be “Please find the degree requirements for the accounting technology associate degree program. Print a print-friendly version of the degree plan.” This question tests “easier to find” by testing the site navigation, link wording, and link placement. It also tests “easier to print” by testing the newly added print-friendly button and the button placement.

3. Create a description of the test and specific instructions to be read to the users who will be testing your site (King, 2003). “Make sure the test description and instructions place emphasis on Web site faults, rather than on the volunteers’ getting correct/incorrect answers on the test” (King, 2003).

4. Select a representative group of users. Select five volunteers who you believe to be representative of your target audience. It is more efficient to test a smaller number of volunteers three times rather than a larger number just one time (King, 2003). Five users are optimal according to Nielsen (2003). It is easier and more affordable to run several small tests and to revise between each one to improve usability flaws as they are encountered.

5. Arrange for a location to run the test. Members of the user group can be tested individually or at the same time, but their work must be independent of one another.

6. Observe but do not assist the users as they navigate your Web site. Take notes of how the users maneuver the site, whether or not they find the answers to your questions, and how long it takes them to find the answers. “Observe what the users do, where they succeed, and where they have difficulties with the interface. Shut up and let the users do the talking” (Nielsen, 2003). However, Nielsen (2003) also warns that listening only to what people say can be misleading, which is why observation is so important to the process.

7. Ask users to complete the instrument in Appendix A. This instrument is adapted from the research of Palmer (2002). Users should complete it immediately upon completing the previous steps to provide the most honest impression of their experience.

8. Study the results of your study. Based on the results, redesign and then do it all over again! These steps will be repeated until your Web site meets your standards for usability.

Step 3: Evaluate Web Design Guidelines

The third and final step in the evaluation process is to determine if a Web site is in compliance with established design guidelines. To avoid violating basic Web design principles, ask your user testing group to complete a Design Guidelines Checklist (Appendix B). The Checklist serves as one last measure to catch any usability/design errors that were not caught in the previous two steps of the evaluation process. The checklist is based on the guidelines published by the Website Usability Testing Center (2001), Nielsen (2003), and Palmer (2002).
DISCUSSION

One of the objectives of anyone producing Web documents is to ensure that the information presented is communicated as intended. In the previous sections, we have presented a review of the literature on readability, usability, and Web design guidelines. In the last section, we presented a three-step evaluation process. Each of these evaluation methods have been used extensively but in many cases independent one another. We are proposing that we use all of what we know to ensure the best quality product. These processes have been proven to improve the readability, usability, and design of Web sites.

For educators, we recommend training students to use this three-step evaluation process in the classroom. Begin by assigning a Web site creation project to groups of students. Then ask students to use the three-step evaluation process to assess the effectiveness of their Web site. This process is an excellent educational tool because it trains students to consider purpose and audience in their design. It also trains students in a real-world approach to Web site testing.

REFERENCES


**Appendices**

Table 1. Interpretation of Flesch Reading Ease Ratings

<table>
<thead>
<tr>
<th>Reading Ease Rating</th>
<th>Difficulty Level</th>
<th>Educational Attainment Level</th>
<th>Typical Style of Magazine</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-30</td>
<td>Very difficult</td>
<td>Postgraduate Degree</td>
<td>Scientific</td>
</tr>
<tr>
<td>30-50</td>
<td>Difficult</td>
<td>Undergraduate Degree</td>
<td>Academic</td>
</tr>
<tr>
<td>50-60</td>
<td>Fairly difficult</td>
<td>Grades 10-12</td>
<td>Quality</td>
</tr>
<tr>
<td>60-70</td>
<td>Standard</td>
<td>Grades 8-9</td>
<td>Digests</td>
</tr>
<tr>
<td>70-80</td>
<td>Fairly Easy</td>
<td>Grade 7</td>
<td>Slick Fiction</td>
</tr>
<tr>
<td>80-90</td>
<td>Easy</td>
<td>Grade 6</td>
<td>Pulp Fiction</td>
</tr>
<tr>
<td>90-100</td>
<td>Very easy</td>
<td>Grade 5</td>
<td>Comics</td>
</tr>
</tbody>
</table>

*Source: Courtis (1986).*

Table 2. Sample Calculation of Readability Statistics

| Readability Statistics | Counts | | Averages | | Readability |
|-------------------------|--------||----------||------------|
| Counts                  |        | |          | |            |
| Words                   |        | |          | | Passive Sentences |
| Characters              | 26     | |          | | Flesch Reading Ease |
| Paragraphs              | 132    | |          | | Flesch-Kincaid Grade Level |
| Sentences               | 2      | |          | |                      |
| Averages                |        | |          | |            |
| Sentences per Paragraph | 3.0    | |          | |            |
| Words per Sentence      | 6.6    | |          | |            |
| Characters per Word     | 4.9    | |          | |            |
| Readability             |        | |          | |            |

*Source: Microsoft Word Readability Statistics*
APPENDIX A – USABILITY SURVEY

WEBSITE: ________________  URL: www.____________

Have you ever visited this website before?  Yes  No

Please circle the response that best reflects your opinion of the Web site you just used.

(1) I find it easy to get this Web site to do what I want it to do.
   Strongly Agree  1  2  3  4  5  Strongly Disagree

(2) The amount of information displayed on the screen was
   Inadequate  1  2  3  4  5  Adequate

(3) The sequence of obtaining information was
   Confusing  1  2  3  4  5  Clear

(4) The information on succeeding links from the initial page was
   Predictable  1  2  3  4  5  Unpredictable

(5) The Web site was
   Frustrating  1  2  3  4  5  Satisfying

(6) The layout of pages made tasks easier
   Never  1  2  3  4  5  Always

(7) The speed in which the computer provided information was
   Fast Enough  1  2  3  4  5  Too Slow

(8) If you had a future need for information presented in this Web site,
   how likely is it that you would consider returning to this site?
   Very Unlikely  1  2  3  4  5  Very Likely

(9) How would you compare this site to other similar Websites?
   Much worse  1  2  3  4  5  Much better

APPENDIX B - DESIGN GUIDELINES CHECKLIST

Response Time

Yes  No

☐ ☐ adheres to 8-second rule (load time is approximately 8 seconds or less)
☐ ☐ progress indicator is included during load time
☐ ☐ images optimized properly

Navigation

Yes  No

☐ ☐ groups buttons, bars, and other aids together
☐ ☐ consistently places buttons, bars, and other navigational aids on each Web page
☐ ☐ uses hyperlink text (not images) accurately to describe the linked pages
☐ ☐ provides a back to home link on every page
☐ ☐ able to distinguish between used and unused links
☐ ☐ able to distinguish between internal and external links
☐ ☐ includes only useful, relevant links (no dead links)
☐ ☐ uses page anchors on large documents/return to top
☐ ☐ if a search engine is included, clearly defines its scope and includes a “Search” button

Structure/Layout

Yes  No

☐ ☐ orders most relevant material at the top of the site’s hierarchical structure
☐ ☐ orders most relevant material at the top of each page
☐ ☐ maintains constant design from page to page
☐ ☐ organizes each page within the site consistently (including navigation)
avoids frames except for navigation
layout is not fixed width (adjusts to user’s screen)
avoids vertical scrolling
offers a number of printing options
uses readable and visually appealing font and format
controls color in an aesthetically pleasing way

Content

Yes No
provides contact information for webmaster and site owner
keeps information accurate and current
tells user when the Web site was last updated
offers up-to-date, relevant FAQs

Notation of contact information on a separate page

Janna B. Arney
1061 Plantation Drive
Brownsville, TX 78526
jbarney@utb.edu