

Background and Methodology

The U.S. Department of Health and Human Services (HHS) Research-Based Web Design & Usability Guidelines (*Guidelines*) project began in March of 2000. Since that time, each guideline presented in this book has undergone an extensive internal and external review. The process used to create the *Guidelines* is presented here.

Step 1: Creating the Initial Set of Guidelines

HHS wanted to develop a set of guidelines that could help designers build Web sites that are based on the best available research. The initial set of guidelines were drawn from existing Web design guideline and style guides, published research articles, research summaries, publicly available usability test reports, and lessons learned from in-house usability tests. This effort resulted in more than 500 guidelines.

Step 2: Reviewing the Initial Set of Guidelines

The initial set of 500 guidelines was far too many for Web site designers to use effectively. An internal review process was conducted to:

- identify and combine duplicate guidelines.
- identify and resolve guidelines that conflicted with each other; and
- reword unclear guidelines.

Each of the reviewers had experience in Web site design, usability engineering, technical communication, software design, computer programming and/or human-computer interaction. This internal review reduced the initial set of guidelines to 398.

Step 3: Determining the 'Relative Importance' of Each Guideline

To determine the 'Relative importance' of each guideline, 16 external reviewers were recruited. Half of these reviewers were Web site designers and half were usability specialists. Each reviewer evaluated each guideline and assigned a rating based on the question, 'How important is this guideline to the success of a Web site?' Those guidelines that were rated as having little importance to the success of a Web site were eliminated. The set of guidelines now was reduced to 287.

Step 4: Determining the 'Strength of Evidence' for Each Guideline

The next step was to generate a 'Strength of Evidence' rating for each guideline. To do this, a group of eight usability researchers, practitioners and authors were recruited. These reviewers were all published researchers with doctoral degrees, experienced peer reviewers, and knowledgeable of experimental design. These reviewers constructed a set of criteria for judging the strength of the evidence for each guideline, which was used as the 'Strength of Evidence' scale.

Step 5: Finding Graphic Examples for the Guidelines

Most of the guidelines required a graphic example to ensure that users clearly understand the meaning of the guideline. The project team identified and reviewed several possible examples for each guideline, and selected the best examples. During this activity, the number of guidelines was further reduced.

Step 6: Grouping, Organizing, and Usability Testing the Guidelines

To ensure that the information about specific Web design issues is easy to find, a group of 20 Web site designers were asked to participate in a formal 'grouping' of the guidelines by participating in a card-sorting exercise. Each of the twenty individuals put the guidelines into groups that reflected how they think about Web design issues, and then provided a name for each group. Data from this exercise was analyzed with specially developed software and formed the chapters of this book.

Several draft page layouts in print format were developed for this book. These drafts were usability tested to determine how best to facilitate readers' ability to locate and understand information on a page. These findings, as well as readers' preferences, served as the basis for the final page layout. The final set that was published in 2004 contained 187 guidelines.

Step 7: Updating the Set of Guidelines

Since publishing the 2004 edition of the *Research-Based Web Design and Usability Guidelines*, the research literature has been continually searched for new and useful research-based information. We identified new relevant research that enabled us to substantially revise (update) 21 existing guidelines, and to add 22 new guidelines. Minor editing changes were made to a few other guidelines. The new and revised guidelines were edited by three different, independent groups of computer professionals. After editing, the final number of guidelines was 209.

The 'Relative Importance' ratings were revised based on a new survey in which 36 Web site professionals responded to an online survey. Each of these people reviewed each of the existing 209 guidelines and rated each one on a Likert-like importance scale with the anchors set at 'Important' to 'Very Important.'

The 'Strength of Evidence' ratings were revised for those guidelines where new research was reported. In this case, 13 usability professionals rated each of the new and revised guidelines, and assigned 'Strength of Evidence' ratings. The raters all were very familiar the research literature, all had conducted their own studies, and there was a high level of agreement in their ratings (Cronbach's $\alpha = .92$). The criteria used for making the 'Strength of Evidence' estimates is shown on the next page.

The 'Strength of Evidence' ratings were revised for those guidelines where new research was reported. In this case, 13 usability professionals rated each of the new and revised guidelines, and assigned 'Strength of Evidence' ratings. The raters all were very familiar the research literature, all had conducted their own studies, and there was a high level of agreement in their ratings (Cronbach's alpha = .92). The criteria used for making the 'Strength of Evidence' estimates is shown below:

5 – Strong Research Support ①②③④⑤

- Cumulative and compelling, supporting research-based evidence
- At least one formal, rigorous study with contextual validity
- No known conflicting research-based findings
- Expert opinion agrees with the research

4 – Moderate Research Support ①②③④○

- Cumulative research-based evidence
- There may or may not be conflicting research-based findings
- Expert opinion
 - Tends to agree with the research, and
 - A consensus seems to be building

3 – Weak Research Support ①②③○○

- Limited research-based evidence
- Conflicting research-based findings may exist - and/or -
- There is mixed agreement of expert opinions

2 – Strong Expert Opinion Support ①②○○○

- No research-based evidence
- Experts tend to agree, although there may not be a consensus
- Multiple supporting expert opinions in textbooks, style guides, etc.
- Generally accepted as a 'best practice' or reflects 'state of practice'

1 – Weak Expert Opinion Support ①○○○○

- No research-based evidence
- Limited or conflicting expert opinion