

**USABILITY STUDY SHOWS MU EXTENSION STYLE GUIDE WEBSITE
IS USABLE AND HIGHLY LEARNABLE**

by

DEEANNA ADKINS

Prof. Jennifer Rowe, Chair

Prof. Jan Colbert

Dr. Sanda Erdelez

MAY 2014

MU Extension has more than 1,200 active employees housed on the University of Missouri campus, other campuses around the state, and in almost every county in the state. We ask these faculty and staff members to use the house editorial style and usage guide to keep all communications, including Web pages and publications, consistent. More than 200 employees have been trained to use the style guide, and more than 200 others have seen demonstrations of the guide. Previously, the style page I developed was on the Web and was simply a long list of all of our styles in an html document that was only searchable using the browser's page search function (e.g., Ctrl+F on a PC).

As we explain to people who are not trained in communications work, style guides are about consistency and efficiency. Having style choices already decided lets the writer and editor concentrate on the content. In my current position as Web coordinator for MU Extension Communications and Marketing, my staff and I coordinate Web pages that are maintained by people all across extension. To maintain some consistency across these, I developed a house style guide for use in MU Extension.

For my master's project, I took that guide, expanded it, and developed a database-driven version that is searchable and customizable. Then I needed answers to some questions: Is the site any better than it was before? Can people use the site in the ways I intended? Have I allowed for users who think differently than I do to easily use the functions of the site? These are questions all communicators, journalists and news organizations should ask about websites and online applications such as this database-driven style guide to be sure users are able to take advantage of the intended functionality.

I planned to evaluate the learnability and usability of the new style guide with the usability testing method, which I have found useful in my work and have been trained to do with a graduate course in usability.

For the research component, I conducted the usability test of the new database-driven version of the MU Extension editorial style and usage guide with three users to measure their experience with the style guide. Krug (2014) stated that three participants is the ideal number for a round of testing that is intended to “improve what you’re building by identifying and fixing usability problems.” All of the users work for MU Extension and use the current style guide for parts of their jobs. The tests were to determine if users had problems finding the information they need using the style guide and if the guide is usable and learnable. I developed nine tasks for participants that allowed me to examine the following research question: How usable and learnable is the navigation and search function of the MU Extension online style and usage guide?

Literature Review

During a major redesign and reorganization of the MU Extension website that started in late 2006, the MU Extension Web team made a conscious decision to base decisions on usability testing with customers and internal staff rather than opinions of the team or other staff. Members of the Web team and a group of researchers from the Information Experience Lab at the University of Missouri School of Information Science and Learning Technologies used heuristic evaluation, focus group interviews and surveys, think-aloud interviewing, and multiple-user simultaneous testing to evaluate the website (Wang et al., 2010). The researchers found that understanding users’ problems,

wants, and needs can change how Web designers make decisions and can make a website more usable and useful for both their customers and internal users.

Purpose and value of a house style guide.

MacKay (1997), in an article that is often cited in later literature on style guides, defined a style guide as “a rule-driven document that sets the parameters for consistency and acceptability for all written materials produced by an individual or group. A house style guide is one that is produced for an organization’s internal use and is specifically tailored for its specific writing contexts” (p. 244). He further noted that there is an assumption that consistent style and form in publications adds to a company’s credibility while inconsistencies detract from the company image.

Almost two decades ago, Allen (1995) suggested four good reasons to develop a corporate, or house, style guide: to create consistency in documents, to promote a professional image, to train newly hired employees, and to define how to generate documents. He also suggested that decreasing costs, which he said would happen as a result of the four reasons above, is the main reason corporations should develop style guides.

Having a house style guide to work from can settle disputes among editors, writers, and subject-matter specialists in an organization (Allen, 1995). Although many answers to style questions can be correct, Bright (2005) stated that “organizations wishing to present a consistent and coherent message must choose one of the correct answers and reject all other options” (p. 42). As MacKay (1997) concluded, “A style

guide's purpose is to provide ground rules, with both the organization and the audience in mind" (p. 250).

The role of the house style guide has changed with the addition of digital content that portrays a company's image. Additional sections are needed for items such as organizational identity guidelines and terminology and other styles specific to the type of media where the document resides (Bright, 2005). Racine (2008) recommended that Web style guides, in particular, should include both editorial and technical standards. For instance, in our case, the way an MU Extension publication is referenced and/or linked is different depending on whether the content resides in a print publication, a Web page to be viewed on a desktop, a Web page to be viewed on a tablet or mobile phone, or a mobile app.

Many of the more recent articles that discuss style guides are talking more about setting technical standards than about editorial standards.

Evaluation of style guides.

Although there are several journal articles about the value of style guides and about the steps for developing effective style guides, there is little concrete research from a user perspective on what works and doesn't work in style guides (MacKay, 1997). McKay said that evaluation of the final product should involve "getting reactions, through a variety of techniques, from users of the style guide" (p. 248) to determine if it is clear, comprehensive, easy to use, attractive, and easy to maintain (Washington, 1993).

Allen (1996) researched user attitudes toward corporate style guides using a survey. His sample included 200 randomly selected attendees of the 40th Society for

Technical Communications conference, so his respondents were primarily writers and editors. Of the 69 respondents who use a corporate style guide, 92.8% said their organization's guide helps them fulfill work responsibilities. Allen concluded that respondents perceived that style guide usage "allows the corporate writer to produce more professional, user-friendly documents in less time without conflict" (p. 238). He stated that his survey reveals benefits of style guide usage such as consistency among documents and time saved on document generation, which validates the reasons given in previous articles that talk about the value of style guides but don't back the conclusions with empirical research.

Usability testing of online databases.

To define usability, Rubin and Chisnell (2008) stated, "When a product or service is truly usable, the user can do what he or she wants to do the way he or she expects to be able to do it, without hindrance, hesitation, or questions" (p. 4). Nielsen (1993) defined usability as multi-dimensional properties of a user interface that are normally associated with the following five attributes: learnability, efficiency, memorability, errors and satisfaction.

I was not able to find usability studies on style guides. However, I was able to find usability tests on some more complex websites with online databases. Many of their usability issues and evaluation techniques are applicable to this project, and the results of two of these studies are discussed below.

The Georgia Tech Library website was first redesigned using information architecture principles for organization then redesigned to accommodate what was

discovered in subsequent usability tests (King and Jannick, 2005). Think-aloud usability tests with inexperienced patrons showed mainly that users didn't know which search interface (e.g., catalog, databases, and e-journals) to use to find what they needed. They also used the Quick Catalog Search like a Google search, so it was removed. The redesign featured ways to guide the user through navigational choices to attempt to reach the correct search.

PENUMAT, which stands for Personal Nutrition Management Tool, is an interactive, Web-based database that includes nutrition management information and screening tools (Bozkurt, et al., 2011). To usability test the database, the researchers used a multi-method approach that included protocol analysis, interviews, and a System Usability Scale survey with a sample of 10 healthy volunteers. Usability problems from think-aloud sessions were sorted by content analysis and grouped into Nielsen's (1993) 10 usability heuristic categories. Each heuristic was found to be violated at least once. The authors concluded that although the SUS scores, which ranged between 77.5 and 100, with a median of 88.7, were acceptable, the multi-method approach was necessary because both the think-aloud sessions and the interviews found usability problems with the website that the SUS scores didn't indicate.

Post-task question.

Directly after a user finishes a task during a usability study can be an opportunity to collect information about the user's experience with that task. Tullis and Albert (2008) listed several different ways to ask a user to evaluate a task that has been completed, including measuring ease of use, using the three-question After-Scenario Questionnaire,

and using an expectation measure that compares how easy or difficult the participant thought the task would be compared to how easy or difficult he or she thought it was going to be before attempting the task.

Tedesco and Tullis (2006) compared five methods, including those above, and determined that measuring ease of use with one simple question (e.g., “This task was easy to complete” with a five-point Likert scale from “strongly agree” to “strongly disagree”) was the most reliable measure, especially with the small sample sizes often found in usability tests.

System Usability Scale (SUS).

The System Usability Scale (Brooke, 1996) was developed as a cost-effective and practical way to test the usability of and user satisfaction with industrial systems. One of its advantages is the ability to compare the scores of multiple systems or to compare successive iterations of one system. Brooke constructed the scale by assembling 50 potential questionnaire items and testing them on two examples of software systems, one considered easy to use and the other considered very difficult to use. Items that elicited the most extreme responses were selected. Half the statements are positive, and half the statements are negative. Brooke suggested that the scale be used after the product has been used by the respondent but before the respondent is engaged in any discussion about the product.

Bangor, Kortum, and Miller (2008) listed four reasons the SUS is a good choice for usability practitioners. First, it is flexible enough to be used for a range of products, websites, and systems. Second, it is fast and easy to use. Third, the SUS score is easy to

understand. And fourth, the survey is free for all to use. Sauro (2011) suggested that data he has analyzed from more than 5,000 SUS surveys across 500 different evaluations have shown that SUS is a reliable and valid measure of perceived usability.

While Brooke (1996) reported that his System Usability Scale measures general usability, Lewis and Sauro (2009) performed a factor analysis on the scale using two independent data sets: a set of theirs that included 324 complete SUS questionnaires and a set of Bangor, Kortum, and Miller's (2008) that included 2,324 SUS questionnaires).

Lewis and Sauro (2009) found a two-factor solution. Items 1, 2, 3, 5, 6, 7, 8, and 9 aligned with a factor they named usability, and items 4 and 10 aligned with a factor they called learnability. They then tested the reliability of the scales. For the overall SUS (all items included), their coefficient alpha was .92 (consistent with Bangor, Kortum, and Miller's finding of a coefficient alpha of .91). For the new usability scale, the coefficient alpha was .91, and for the new learnability scale, it was .70. They concluded that all scales met the minimum standard of .70; therefore, the SUS could be used to determine scores for overall usability, usability, and learnability. They considered the eight-question usability score to be a cleaner estimate of usability than the overall usability score but recommended keeping the two learnability items in the scale for the additional measure of learnability.

Learnability of websites.

Nielsen (1993) called learnability of a system the most fundamental usability attribute because in general, a system needs to be easy to learn and because learning to use the system is usually the user's first experience with it. Rubin and Chisnell (2008) stated that learnability can also refer to a user's ability to relearn a system that he or she uses infrequently.

Lewis and Sauro found that two items in the System Usability Scale, 4 ("I think that I would need the support of a technical person to be able to use this website.") and 10 ("I needed to learn a lot of things before I could get going with this website."), aligned with a factor they called "learnability" because the commonality in the two items is about the ease or difficulty of learning to use the website.

Theoretical Framework

In the field of human-computer interaction, which includes interaction design and interface design, prescriptive theories are considered to be guidelines for system design (Knudtzon, 2002). Sharp, Rogers, and Preece (2007) referred to these guidelines as frameworks that offer advice to designers as they develop and design the user experience. While they said the frameworks are generally based on human behavior theories, they commonly are also formed from the results of user studies and other findings in design evaluation.

Shneiderman and Plaisant (2004) referred to these prescriptive theories as principles that guide the design of the human-computer interface and considered them

more widely applicable and enduring than, for example, guidelines documents developed by software development companies.

One set of prescriptive theories, or principles, that designers find useful when developing human-computer interfaces are Nielsen's (1993) 10 usability heuristics.

Nielsen's 10 heuristics.

Nielsen's (1994, p. 30) 10 usability heuristics, which I used as guidelines when developing the interface of the database-driven version of the MU Extension style and usage guide, are listed below with a brief summary of each heuristic. These heuristics are revised from Nielsen's 1993 version after a factor analysis of 249 usability problems. Although they were written almost two decades ago, these heuristics are still used as a standard for design development and heuristic evaluation of interfaces as shown by the study of a Web-based nutrition database conducted by Bozkurt, et al., in 2011. I also list which heuristics were met and which were violated based on the results of the study in the Conclusion and Recommendations section.

1. "Visibility of system status." Users should always know what is going on in the system through appropriate feedback.

2. "Match between system and the real world." Terminology used in the system should be familiar to the user rather than system-oriented. Information should be presented in a natural, logical order.

3. "User control and freedom." Wherever the user ends up in the system, the way back to the previous state should be clearly visible.

4. “Consistency and standards.” Information should be presented consistently so users don’t have to wonder if different terms or actions mean the same thing, and industry standards should be followed when possible.

5. “Error prevention.” Use careful design to keep errors from happening.

6. “Recognition rather than recall.” Users should not be expected to remember things from one part of the site to another. Help and instructions should be available whenever needed.

7. “Flexibility and efficiency of use.” The system should be designed for efficient use by both expert and novice users.

8. “Aesthetic and minimalist design.” Designers should include only relevant information in the interface design.

9. “Help users recognize, diagnose, and recover from errors.” Error messages should be precise, clear, and polite and help the user solve the problem.

10. “Help and documentation.” While a system is better if users don’t need documentation, any necessary help information should be easy to search, be focused on users’ tasks, be concise, and list concrete steps for the user to follow.

Method

Purpose and objectives.

I designed the new database-driven interface using the 10 Nielsen heuristics listed above with the assistance of project design prototypes made while taking a Coursera.org course on Human-Computer Interaction completed while doing my project work. After this, I evaluated the usability and learnability of the new database-driven style and usage

guide website with three users in usability tests with tasks written to approximate the breadth of what they can do using the interface. The usability tests included observation of the users as they perform the tasks, a one-question, post-scenario survey after each task (Tedesco and Tullis, 2006; Tullis and Albert, 2008), and a post-test System Usability Scale (SUS) (Brooke, 1996) survey (modified for websites) to evaluate the style guide's overall usability, usability and learnability (Lewis and Sauro, 2009).

Research question.

I examined the following research question: How usable and learnable is the navigation and search function of the MU Extension online style and usage guide?

Participants.

I selected participants from a pool of MU Extension employees who have been introduced to the current style page and asked to use it for their work. One was in a northeast Missouri county, and the other two work on the MU campus. The three participants have the following demographic characteristics and Web and style guide usage levels:

- Age: One between 20 and 29; two between 50 and 59
- Gender: All female
- Highest education level completed: One, some college; two, master's degrees
- Comfort with online technology: All use the Web every day
- General style guide usage: Two use a style guide every day; one uses a style guide once a month

- MU Extension Style Guide usage: One uses the guide every other day; one uses it once a week; one uses it once a month

Test environment.

The studies took place in March and April 2013. I conducted two studies in the University of Missouri Agriculture Building, in the offices of MU Extension and Agricultural Information. I conducted a third study in a county office in northeast Missouri. Participants used a Windows PC laptop running the Firefox browser with a high-speed connection to the Internet and a Web camera and microphone attached. The Web camera captured the participants' faces, the microphone captured the participants' voices, and the Morae software recorded what happened on the screen. I had some issues with audio quality during two of the sessions. However, while the participants were asked to think aloud, they made few comments. I had captured the meaningful comments in my notes as I sat next to them during the tests, and I was able to easily match them to the video in Morae.

I conducted three individual usability test sessions. Before the test, participants reviewed and signed a video release form. I used a few minutes at the beginning of each session to explain the testing process to the participant and a few minutes at the end of each session for a post-test debriefing interview. I acted as the moderator during the tests and watched the recordings after to match my notes to the recordings using markers in Morae. During the middle of the session, participants performed the nine tasks found in Table 1. I used Morae's autopilot feature to present a demographic questionnaire at the start of the test, the post-task question after each task, and a System Usability Scale

survey (see Appendix A for the survey questions) at the end as well as to log the start and end of tasks.

Table 1
Tasks used in this usability test and the criteria for successful completion of each task

Task	Description	Success criteria
1	You will need to sign in to the style guide with your pawprint and password to use all of the guide’s features. Can you find how to sign in?	Find the login in button, choose the correct domain, and enter username and password.
2	One of your Web pages has links to several specialists’ email addresses. Sometimes the link is on the person’s name, and sometimes the link is after the name on the written-out email address. You want to change these to all be consistent and to follow style. How would you look for MU Extension’s style in this case?	Find style entry called “contact information.”
3	When editing your documents for the MU Extension style guide, you can never remember if the correct spelling is Web site, website or web site. How would you find MU Extension’s style for this?	Find style entry called “website.”
4	The style guide includes examples for some styles, but you keep running across a date format that isn’t listed in the examples. Is there a way you can add this example to the style guide so you can remember it in the future?	Must be logged in. Use “Add a comment to this style” link (found under the term of each style) to add an example in the comments.
5	If you would like to ask someone to add the date format example to the style guide so everyone, not just you, can see it, how would you proceed?	Use the mailto: link at the top that says “Comments or questions? Email the editors.”
6	You are writing a story about Jane Doe. After you’ve mentioned her full name	Find the style entry called “courtesy titles and name on

Task	Description	Success criteria
	once, what do you call her throughout the rest of the article? Jane, Jane Doe, Doe, Ms. Doe? How would you find out what to call her in the style guide?	second reference.”
7	You are writing about the Livestock Symposium that is held in Kirksville every year. You are never sure when the word Missouri should be abbreviated or how it should be punctuated. How would you find MU Extension’s style for this?	Find the style entry called “state names.”
8	The editors have noticed that many extension materials include the same style mistakes, so they made a way to view often-misused styles. Can you find how to view the list?	Choose “Show advanced search options” then check “Frequently needed” under “Show styles that are tagged.”
9	Although the editors can mark what they consider frequently needed styles, there are other styles you often need to look up. Is there a way for you to keep track of your own frequently needed styles?	Must be logged in. Use “Add to My selections” link (found under the term of each style) to add to the user’s list of frequently needed styles.

The SUS survey I used in this study is modified slightly from Brooke’s (1996) original survey with the recommended adjustments from Bangor, Kortum, and Miller (2008) and Tullis and Stetson (2004). Specifically, throughout the survey, the word “system” was changed to “website” in each statement to better describe the Web-based interface (Tullis and Stetson, 2004). Statement 5 was modified from “I found the various functions in this website were well integrated” to “I found the various search and sorting functions in this website were well integrated” to fit this particular style and usage guide website based on Jarrett’s (2011) suggestion to minimize user questions by modifying the statements to fit the interface being tested. In statement 8, the term “cumbersome” was

changed to “awkward” based on Bangor, Kortum, and Miller’s (2008) experience with respondents’ lack of understanding of the word cumbersome.

Data collected.

I collected performance and preference data using markers in Morae and notes from observing the sessions as well as answers to demographic questions, a post-task question and an overall System Usability Scale survey that was modified for websites.

Performance data.

- Time to complete each task (time on task)
- Count of incorrect navigation choices
- Number of times participant was prompted
- Number of tasks completed with and without assistance

Preference data:

- Ease of use overall (measured with system usability scale, called SUS)
- Learnability (measured using SUS items 1, 2, 3, 5, 6, 7, 8, and 9)
- Usability (measured using SUS items 4 and 10)
- Ease of task completion (measured with post-task question)

Results and Discussion of Findings

Overall, the qualitative and quantitative performance and preference results show that tasks 4 and 8 were the most difficult for participants to complete. Task 2 was also difficult for the participant who was least familiar with style guides.

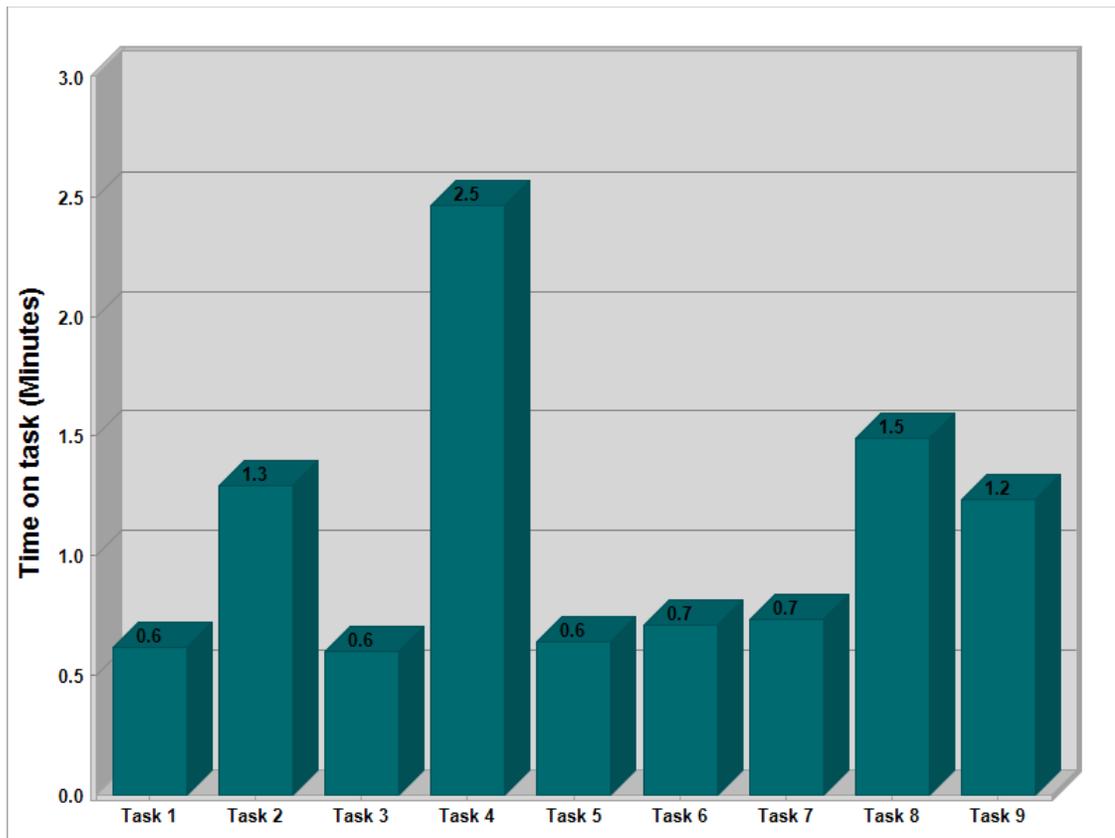
Performance data.

Time to complete each task.

Time to complete all tasks was generally less than a minute and a half as shown in

Figure 1. All participants completed five of the tasks in less than a minute per task.

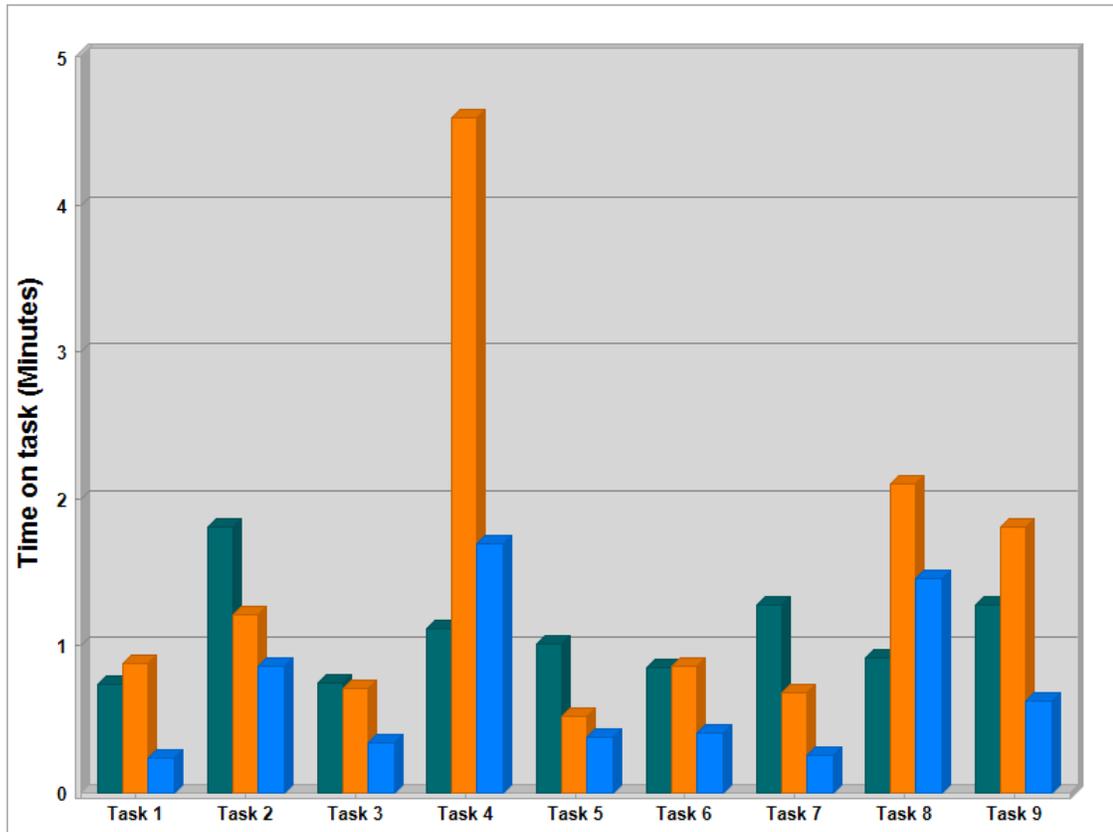
Figure 1
Average time on task by task



Task 4, however, required four and a half minutes for one participant as shown in Figure 2. She thought clicking the “Styles personalized for me: My selections” checkbox completed the task when she actually needed to find the function “Add a comment to this style” to complete the task. It took her awhile to learn how to clear her selections using

the “Clear all” button, which led to the recommendation to make that button say “Clear all selections.”

Figure 2
Time on task, by participant



Because I was using autopilot in Morae, one participant, represented by the teal bar in Figure 2, accidentally “ended” task 8 early, so her time on task is not accurate for task 8.

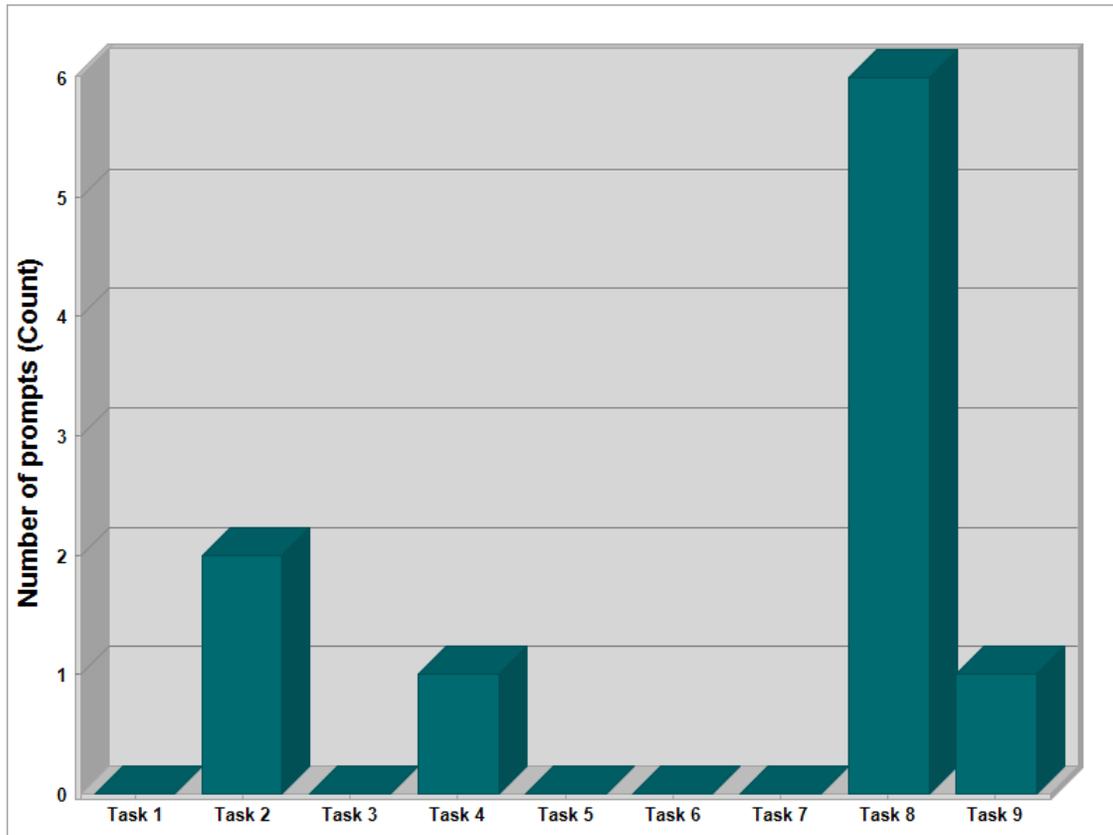
Count of incorrect navigation choices. An incorrect navigation choice in this study is defined as when the participant thought she had finished the task but had not found the correct style or database function that completed the task. Two participants made incorrect navigation choices, one in task 2 and one in task 4. Both times, the

participants were prompted to look somewhere else and then found the desired style or function. The participant who made the incorrect navigation choice in task 2 thought she found the correct way to style specialists' email addresses under the "contact information" entry when the information was actually under the "email addresses on the Web."

The participant who made an incorrect navigation choice in task 4 thought clicking the "Styles personalized for me: My selections" checkbox completed the task when she actually needed to find the function "Add a comment to this style" to complete the task.

Number of times participant was prompted by the moderator. Figure 3 shows participants were prompted most in task 8 (each participant was prompted twice during this task). Two issues arose during this task. First, participants overlooked the link that said "Show advanced search options." Second, after watching the participants attempt to complete the task, I believe the difficulty in this task was exacerbated by the verbiage in the task. The task read: "The editors have noticed that many extension materials include the same style mistakes, so they made a way to view often-misused styles. Can you find how to view the list?" A challenge in writing tasks is trying not to lead a participant directly to the thing he or she is trying to find by using the actual words on the screen. However, this can backfire when a participant decides she must find exactly the words used in the task. The participants were looking for "often-misused styles" which didn't easily translate in their minds to "frequently needed," which is the terminology of the database tag. Also, there was nothing to indicate this might be an advanced search option.

Figure 3
Number of times a participant was prompted by the moderator

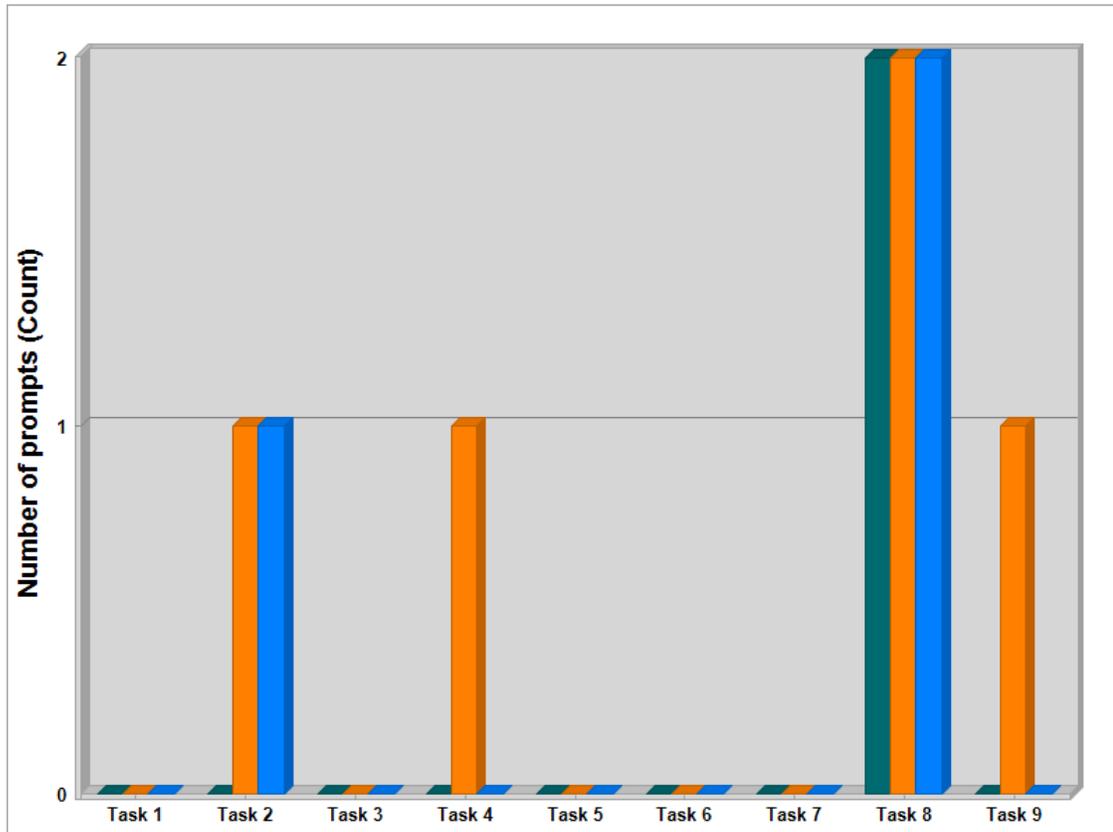


Two different participants were prompted during task 2, as shown in Figure 4. One participant asked questions about the meaning of the task. I prompted the other participant after she thought she found the correct way to style specialists' email addresses under the "contact information" entry when the information was actually under the "email addresses on the Web." I asked her if she might look somewhere else for the information.

During task 4, one participant thought she had completed the task, and I prompted her by asking if she was sure she had found the database function that would allow her to add a comment to a style. The same participant was prompted during task 9 when she

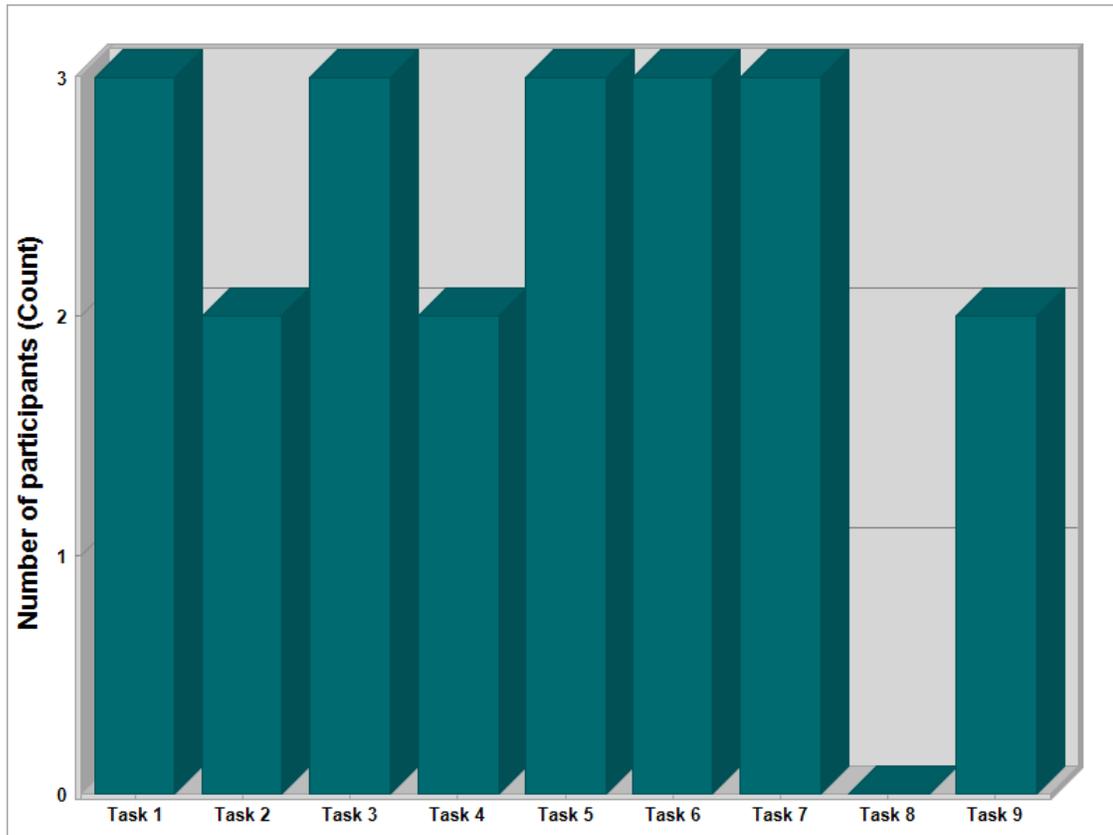
didn't clear her selections. As mentioned before, the confusion during task 9 contributed to the recommendation to change the "Clear all" button to say "Clear all selections."

Figure 4
Number of times a participant was prompted by the moderator, by participant



Number of tasks completed with and without assistance. The number of tasks completed without assistance directly relates to the previous section, "Number of times participant was prompted." Figure 5 shows that five of the nine tasks, tasks 1, 3, 5, 6, and 7, were completed by all participants without prompting. Only one participant needed assistance for tasks 2, 4, and 9. All participants needed assistance twice during task 8.

Figure 5
Tasks completed without assistance



Preference data.

Ease of task completion. After each task, participants were asked to respond to this statement: “This task was easy to complete.” As Table 2 shows, all participants strongly agreed that tasks 1, 6, and 7 were easy to complete and agreed that tasks 5 and 9 were easy to complete. No participant strongly disagreed or disagreed with the statement for any of the tasks. Only one participant each marked tasks 4 and 8 at the mid-point between strongly agree and strongly disagree. This corresponds with previous data that showed tasks 4 and 8 as the most difficult to complete correctly and without assistance.

Table 2

Number of participant responses to the post-task question “This task was easy to complete” for each task

Task	Strongly agree				Strongly disagree
1	3	-	-	-	-
2	1	2	-	-	-
3	2	1	-	-	-
4	1	1	1	-	-
5	-	3	-	-	-
6	3	-	-	-	-
7	3	-	-	-	-
8 ¹	-	1	1	-	-
9	-	3	-	-	-

¹One participant accidentally “ended” task 8 early and did not answer the post-task question.

Ease of use overall. After completing the nine tasks, participants were asked to take the 10-question SUS survey, which included an open-ended comment box. The overall average SUS score for the site was 78.33; the highest score possible is 100. Tullis and Albert (2008) equated the scores to percentages for ease of interpretation, and they said an average SUS score under 60 percent is relatively poor while a SUS score of more than 80 percent is considered pretty good. Sauro (2011) conversely said SUS scores are not percentages and used a process called normalizing to convert SUS scores to percentages and, subsequently, letter grades that are easy to use to describe a website’s score. He said a SUS score of 68 is considered “average,” and according to his blog, the score of 78 in this usability test would normalize to a letter grade of B+. The 78 would also be close to a high, “pretty good” score according to Tullis and Albert.

Only one participant chose to use the comment box, and she wrote: “Overall, this website was intuitive and easy to use.”

Learnability and usability factors. Lewis and Sauro (2009) furthered divided the SUS survey into two scales: learnability and usability. Calculated using their formula, the average SUS usability score of the style and usage guide is 75. The SUS learnability score is 91.67. Thus, using Sauro’s (2011) normalized letter grades, the style and usage database as presented to these participants scored a letter grade of B for usability and an A+ for learnability. This was encouraging, and these scores should only increase with improvements to the interface.

Conclusion and Recommendations

The MU Extension style and usage website was designed and evaluated using Jakob Nielsen’s (1994, p. 30) 10 usability heuristics, which are listed in the Theoretical Framework section above, as guidelines. I will match which Nielsen heuristics are met or violated with the findings from the usability study. Table 3 shows how many times I considered each heuristic to be met or violated based on the usability results. Specific recommendations to improve the site follow each of the findings.

Table 3

Number of Jakob Nielsen's 10 usability heuristics considered met and violated during the evaluation

Heuristic	Number of times heuristic was met	Number of times heuristic was violated
Visibility of system status	1	3
Match between system and the real world	1	2
User control and freedom	1	4
Consistency and standards	3	0
Error prevention	0	2
Recognition rather than recall	1	0
Flexibility and efficiency of use	2	0
Aesthetic and minimalist design	0	1
Help users recognize, diagnose, and recover from errors	0	2
Help and documentation	0	1

Positive findings.

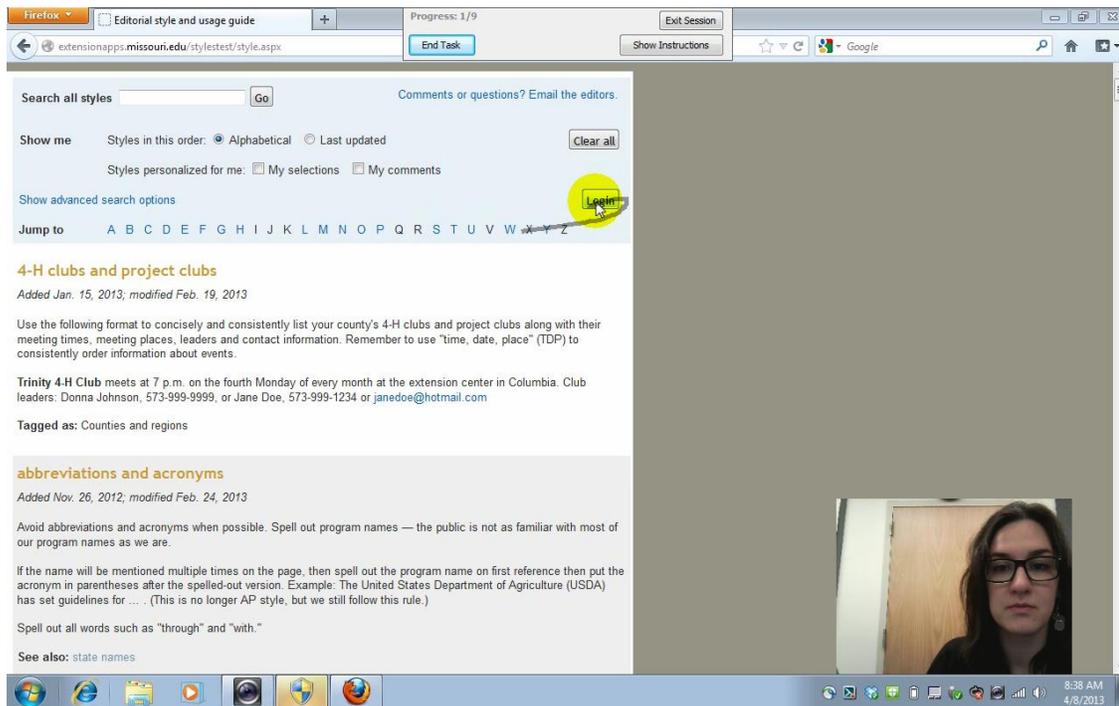
1. The style guide had a very high score for learnability (A+), a B+ for overall usability, and a B for usability. Any improvements should increase those numbers, which are already “good” grades.
2. Participants were able to complete most of the tasks without assistance.
3. Participants liked the ability to search the database rather than always needing to scroll through the page. One participant commented, “It’s interesting that I’m using the search function every time and not clicking down here.” She was pointing to the “Jump to” alphabetical line as she made the comment.

Nielsen heuristics met. Flexibility and efficiency of use; Consistency and standards

4. Participants could easily see how to log in to the database. Figure 6 shows one participant starting the login process. She saw the “Login” button within a few seconds of starting the task.

Nielsen heuristics met. Visibility of system status; Match between system and the real world; User control and freedom; Consistency and standards; Recognition rather than recall

Figure 6
A participant successfully finds the “Login” button to complete task 1

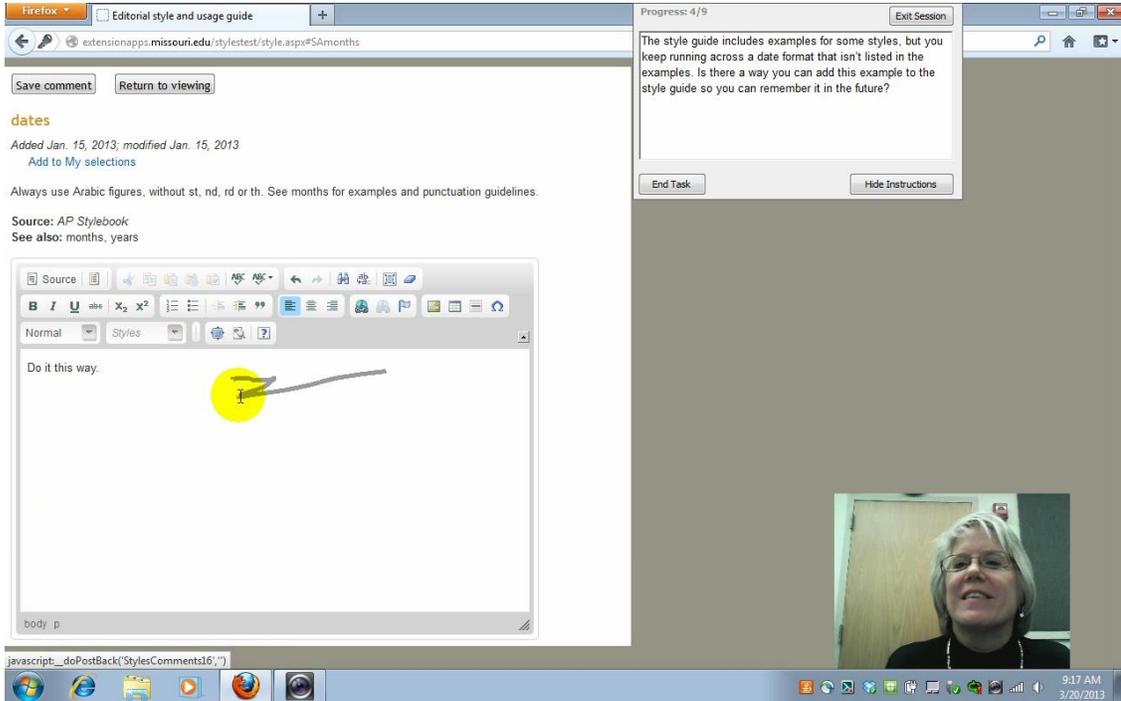


5. Participants, especially the two who were more familiar with style guides, liked the personalization features. One participant said about the ability to mark styles as “My selections” in the database: “That would be a really good way for a ... quick reference, I guess, to the things that I can’t remember.” Another, pictured in Figure 7, when she was able to add a comment to the database, said, “That’s cool.”

Nielsen heuristics met. Flexibility and efficiency of use; Consistency and standards

Figure 7

A participant successfully added a comment to the database during task 4



Users' difficulties and recommended solutions.

The potential solutions to the following difficulties would provide the greatest improvement to the function of the new database-driven MU Extension editorial style and usage guide based on this usability study.

1. It is hard to scroll back to the top of the database after using the “Jump to” alphabetical links. As she scrolled to the top of the page, one participant said, “The only way to get to the top is to scroll.”

Nielsen heuristics met. User control and freedom

Possible solution. Add links to each style that return the user to the top.

2. The advanced search options link is not obvious to users.

Nielsen heuristics met. Help and documentation; Aesthetic and minimalist design

Possible solution. After considerable thought and some discussion with other editors, I decided to recommend leaving this alone rather than putting the advanced information on the main screen because the advanced search options are ones only potential “power users” might use often. Also, the task to find this, task 8, had some issues with verbiage and understandability. Therefore, in keeping with the Nielsen heuristic “Aesthetic and minimalist design,” I recommend keeping the options under the advanced search link off of the main screen. Because the style guide was shown to be highly learnable by the learnability SUS score, the advanced features could be shown in a small user guide on the Web and during style trainings.

3. The “Show styles that apply to all uses” button under the advanced search options didn’t make sense to users.

Nielsen heuristics met. Match between system and the real world

Possible solution. The idea behind this was to show styles that applied to all media. Upon further thought, I recommend removing the wording altogether as it is more difficult on the data-entry end than I previously thought to decide when a style applies to only one medium.

4. “Clear all” button is not obvious to users.

Nielsen heuristics met. Visibility of system status; User control and freedom

Possible solution. Make the button say “Clear all selections.”

5. Add a way to know you have to log in to use the “My selections” and “My comments” features.

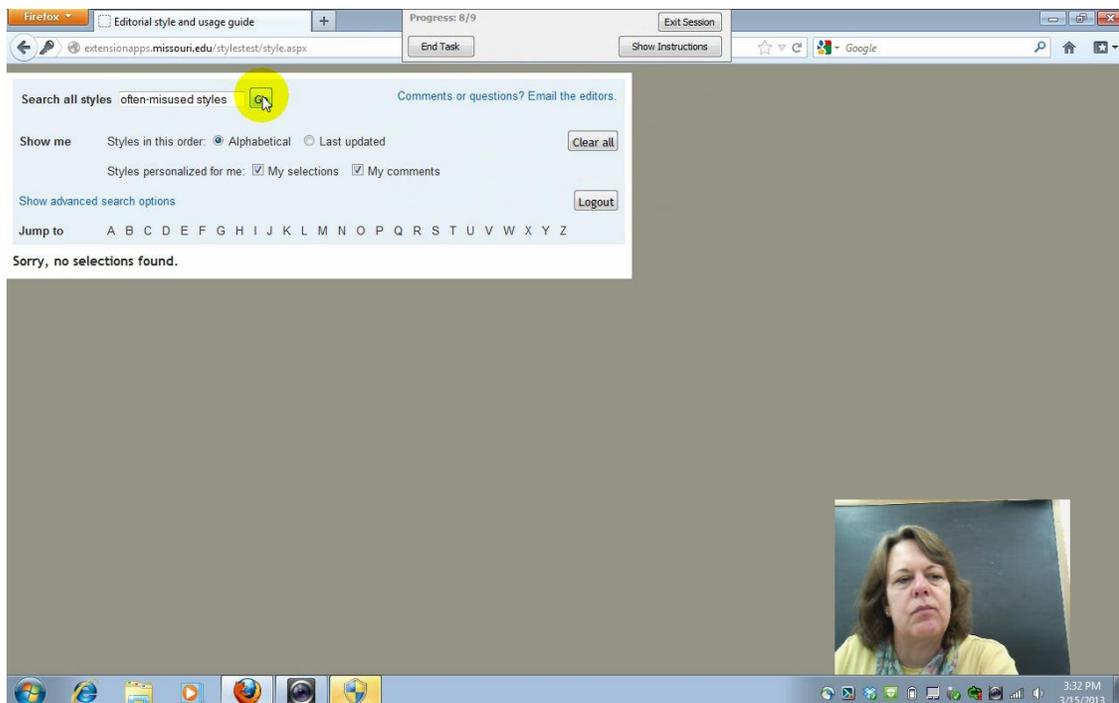
Nielsen heuristics met. Visibility of system status; User control and freedom; Error prevention; Help users recognize, diagnose, and recover from errors

Possible solutions. Add a parenthetical statement under the personalization line that says “(You must log in to use personalization feature.)” and/or add information to the error message that says they must be logged in to use the features.

6. Add a way to know that when you are searching by “My selections” or “My comments,” you must have previously marked some styles as your selections or added comments to a style. In Figure 8, a participant is frustrated when checking both the “My selections” and “My comments” boxes yielded no results even though she had not previously made any selections or added any comments.

Figure 8

A participant is confused when checking both “My selections” and “My comments” yields no results during task 8



Nielsen heuristics met. Visibility of system status; User control and freedom; Error prevention; Help users recognize, diagnose, and recover from errors

Possible solution. Make the error message clear that you must add a selection or comment to be able to search using those attributes.

7. Users with less style guide experience don't realize that when they use a style guide, just like when they use a dictionary, often the term itself is the indication of how the term should be used or spelled. For instance, the "website" entry didn't have further details in it because the term "website" at the top of the entry was meant to indicate that website is spelled as one word. This is a data issue rather than a database or interface design issue and can easily be addressed.

Nielsen heuristics met. Match between system and the real world

Possible solutions. Add the words "always used as one word" in the style description or write out something such as "The word website is always one word and lowercased."

Implications

This study was a somewhat unique combination of using heuristics, or the eye of the expert, to guide the initial database development and usability testing, or the eye of the user, to evaluate the product and implement changes based on the results. Usually, only one of these two methods is used on a product. In this case, the professional project component was to build a practical application and, as a journalist, I used my knowledge of style guides and editing as well as Nielsen's heuristics to design a database application that made sense to me and fit the organization.

Every time I have conducted or observed usability testing, I have found both that some things are easier to do on the site than I would have thought and that other things are not as intuitive as I had imagined. This research was no exception. It is easy to add your own bias and knowledge to the design of a database or website.

After designing the database, I needed to evaluate how others could navigate the database as many of our employees are not trained editors or journalists. The usability testing for the research component added sufficiently rigorous, somewhat ethnographic, research for efficient practical evaluation of the editorial style and usage guide.

In this study, the finding that was most surprising to me was when a user didn't realize the term as written defined how it was to be used. From my perspective and training in use of style guides, that was something I never would have considered to be an issue. However, many members of the intended audience for this guide might have never used a style guide before. It is important to keep in mind that a small number of users can show many usability successes and failures.

The goal of news and informational websites, much like the MU Extension site, is to reach users and allow them to use the functionality of the site to get information and interact with the organization. To be helpful, usability evaluation does not have to involve a large study with many participants. Any organization can add the "human-size" usability testing described here to determine if a site that has changed is improved and more usable or learnable and if users can use the functions and features as the developer intended.

References

- Allen, P. R. (1995). Save money with a corporate style guide. *Technical Communication*, 42, 284-289.
- Bright, M. R. (2005). Creating, implementing, and maintaining corporate style guides in an age of technology. *Technical Communication*, 52, 42-51.
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24, 574-594.
- Bozkurt, S., Zayim, N., Gulkesen, K. H., Samur, M. K., Karaagaoglu, N., & Saka, O. (2011). Usability of a web-based personal nutrition management tool. *Informatics for Health and Social Care*, 36, 190-205.
- Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry*. London: Taylor & Francis.
- Jarrett, C. (2011, February 3). SUS: a good enough usability questionnaire [Web log message]. Retrieved from http://www.rosenfeldmedia.com/books/survey-design/blog/sus_a_good_enough_usability_qu/
- King, H. J., & Jannick, C. M. (2005). Redesigning for usability: Information architecture and usability testing for Georgia Tec Library's website. *OCLC Systems & Services*, 21, 235-243.
- Knudtzon, K. (2002). *Theories in computer human interaction: Prescriptive theories*. Retrieved March 17, 2012, from <http://www.cs.umd.edu/class/fall2002/cmsc838s/tichi/prescriptive.html>
- Krug, S. (2014). *Don't make me think, revisited: A common sense approach to Web and mobile usability*. New Riders.
- Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. *Proceedings of the Human Computer Interaction International Conference (HCII 2009)*. San Diego CA, USA.
- MacKay, P. D. (1997). Establishing a corporate style guide: A bibliographic essay. *Technical Communication*, 44, 244-251.
- Nielsen, J. (1993). *Usability engineering*. San Francisco: Morgan Kaufmann Publishers Inc.
- Nielsen, J., & Mack, R. L. (1994). *Usability inspection methods*. New York: Wiley.

- Racine, J. (2008). Creating a college Web style guide: Defining value, process, and content (Master's thesis). Retrieved from <http://www.racinewebworks.com/RacineThesisFinal.pdf>
- Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests* (2nd ed.). Wiley.
- Sharp, H., Rogers, Y., & Preece, J. (2007). *Interaction design: Beyond human-computer interaction* (2nd ed.). Wiley.
- Shneiderman, B., & Plaisant, C. (2010). *Designing the user interface: Strategies for effective human-computer interaction* (5th ed.). Boston: Addison-Wesley.
- Tedesco, D. P., & Tullis, T. S. (2006). A comparison of methods for eliciting post-task subjective ratings in usability testing. *Usability Professionals Association Conference*, June 12-16, Broomfield, CO.
- Tullis, T. S., & Albert, B. (2008). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*. Amsterdam: Elsevier/Morgan Kaufmann.
- Tullis, T. S., and Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. *Usability Professional Association Conference*, June 7-11, Minneapolis. Retrieved March 29, 2012, from <http://home.comcast.net/%7Etomtullis/publications/UPA2004TullisStetson.pdf>
- Wang, X. C., Yadamsuren, B., Paul, A., Adkins, D., Laur, G., Tawfik, A., & Erdelez, S. (2010). Iterative usability evaluation for an online educational web portal. *International Journal of Multimedia Data Engineering and Management (IJMDEM)*, 1(4), 31-49. doi:10.4018/jmdem.2010100103
- Washington, D. A. (1993). Creating the corporate style guide: Process and product. *Technical Communication*, 40, 505-510.

Appendix A

Pre-study demographics questionnaire

1. What is your age?

19-29
30-39
40-49
50-59
60-69
70-79

2. What is your gender?

Male
Female
Prefer not to answer

3. What is the highest level of education you have completed?

High school
Some college
2-year college degree
4-year college degree
Some graduate work
Master's degree
Professional degree
Doctoral degree

4. Which of the following best characterizes your Web usage?

I use the Web every day.
I use the Web every other day.
I use the Web once a week.
I use the Web once a month.
I never use the Web.

5. Which of the following best characterizes your use of all style guides, both printed copies and online guides?

I use a style guide every day.
I use a style guide every other day.
I use a style guide once a week.
I use a style guide once a month.
I never use a style guide.

6. Which of the following best characterizes your use of the MU Extension Editorial Style and Usage Guide, which is found on the MU Extension website?

I use the style guide every day.
I use the style guide every other day.
I use the style guide once a week.
I use the style guide once a month.
I never use the style guide.

Post-task question (asked after each of the nine tasks)

- Scale for all questions: Strongly agree 1 2 3 4 5 Strongly disagree

This task was easy to complete.

Post-study questionnaire (System Usability Scale)

- Scale for all questions: Strongly agree 1 2 3 4 5 Strongly disagree

I think that I would like to use this website frequently.
I found the website unnecessarily complex.
I thought the website was easy to use.
I think that I would need the support of a technical person to be able to use this website.
I found the searching and sorting functions in this website were well-integrated.
I thought there was too much inconsistency in this website.
I would imagine that most people would learn to use this website very quickly.
I found the website awkward to use.
I felt confident using the website.
I needed to learn a lot of things before I could get going with this website.

Please add any comments or suggestions you have for the MU Extension Editorial Style and Usage Guide. (open-ended with comment box)