The usage of interface design guidelines

Steve Szigeti

August 20, 2007

Abstract

The present study collected data on the attitude of human factors engineering students towards the application of guidelines to their design work. Eighteen participants, assigned by the class instructor to design a Wiki page, were recruited from an undergraduate Mechanical and Industrial Enginerring class at the University of Toronto. A twenty-five guideline collection, drawn from the *Research Based Web Design & Usability Guidelines* (Health and Human Services, 2006), was distributed to participants. They were instructed to apply any particular guideline which they felt was applicable to their design, while noting guidelines which were either not applicable, or unclear. A questionnaire and a focus group session were employed to collect data. It was found that the participants used a combination of experiential knowledge and the searching out of similar sites to design the Wiki pages. Guidelines were viewed favourably, yet the role they play in the design process was found to be minimal.

1.0 Introduction

Design guidelines offer those involved in the design process, both at the novice and expert level, access to expert opinion, research-based evidence or both. From the broad collection of Smith and Mosier's *Guidelines for Designing User Interface Software* (1986) to the tighter focus of the recently published *Research-Based Web Design & Usability Guidelines* (U.S. Department of Health and

Human Services [HHS], 2006), guidelines represent a means to codify both the experience of communities of practice and published research. Good design is difficult, but guidelines may offer assistance, a potential means to positively influence the effectiveness of a website.

In just over 20 years, the literature in interface design guidelines has moved from the need to compile a comprehensive guideline collection (Mosier & Smith, 1986) to questioning whether the use of guidelines can lead to improved user performance (Cherry, Muter, & Szigeti, 2006). The literature has numerous examples of research considering problems with how the guidelines are understood and applied by those involved in the design process (Ivory & Megraw, 2005; Ivory, Sinha & Hearst, 2001; Vanderdonckt et al., 2000; Tetzlaff & Schwartz, 1991), tools that would better facilitate this process (Sutcliffe, Kurniawan, & Shin 2005; Beier & Vaughan, 2003; Henninger, Lu, & Faith, 1997; Henninger, Haynes, & Reith, 1995; Ianella, 1995; Cohen *et al.*, 1994), and finally research that considers the effect of guideline compliance on user performance (Cherry, Muter, & Szigeti, 2006; Ray, 2002; D'Angelo & Twining, 2000; Cherry, 1998). While the relative volume of research considering how best to use guidelines in the design process is high, the literature does not lead to conclusive findings.

Whether design guidelines would assist in the design of computer interfaces was a question posed in the introduction to early guideline sets (Smith and Mosier, 1986). Research followed which considered how the guidelines are understood and applied by those involved in the design process (Ivory & Megraw, 2005; Ivory, Sinha & Hearst, 2001; Vanderdonckt *et al.*, 2000; Tetzlaff & Schwartz, 1991). There remains no clear answer in the literature.

The present study collected data on the attitude of human factors engineering students towards the application of guidelines to their design work. How designers react towards the use of guidelines has not been extensively studied. The author is aware of four papers which discussed the attitudes of designers. Tetzlaff and Schwartz (1991) report that the interpretation of guidelines

appeared to be a significant problem and they suggested the improvement of the presentation of guidelines through the extensive use of examples. Regarding the understanding of guidelines by designers, it has been suggested that the use of jargon or discipline-specific vocabulary is a source of problems (Vanderdonckt *et al.*, 2000), that a lack of specificity within the guidelines makes interpretation difficult (Ivory & Megraw, 2005), and similarly, that the "high level" generalization of guidelines makes operationalizing them challenging (Ivory, Sinha, & Hearst, 2001).

2.0 Method

2.1 Participants

Eighteen participants were recruited from the Ergonomic Design of Information Systems class (MIE 344) at the University of Toronto.

2.2 Materials

The most comprehensive collection of evidence-based guidelines was published in 2006 by the U.S. Department of Health and Human Services. This collection built on an earlier edition (2003), which in turn was developed from a guideline site published by the National Cancer Institute of the United States (2001). The authors of these publications considered both expert opinion and evidence-based guidelines in an attempt to compile a single source for those involved in the design process. Comprised of 209 guidelines divided into 18 categories, the 2006 collection rates the level of research support for each guideline using a five bullet scale. For the purposes of the present study, the three categories of guidelines which contained the largest number of guidelines with strong research support (a rating of five bullets); Layout, Text and Images were chosen. A total of twenty-five guidelines were used.

2.3 Design

Students in the MIE344 course were tasked, in groups of three, by the course instructor to design a Wiki page. Wiki refers to server software that allows users to create and edit Web page content via an internet browser (Wiki, 2002). Most Wiki pages use a design template within which the user can make limited modifications. The Wiki environment can be quite restrictive in terms of design, allowing users freedom to edit textual content, but not the visual design of the page itself. This limitation resulted in the inapplicability of some of the guidelines in this study, although there was disagreement among the participants regarding exactly which guidelines these were (see *Results* below for further discussion.)

The collection of guidelines was distributed to the participants. They were asked to consider each guideline in the context of their design work. As well, they were instructed to apply any particular guideline which they felt was applicable to their design. Guidelines which were either not applicable, or unclear, were to be noted.

After four weeks, a questionnaire was distributed to the participants. They were asked to indicate which guidelines they used, which guidelines they did not apply to their design and/or which guidelines were only partially used. Finally, two focus groups were run by the author lasting just under an hour each during which the experience of working with the guidelines was explored.

3.0 Results

3.1 Questionnaire

Participants were tasked with completing a questionnaire asking them to both report on the usage of each of the twenty-five guidelines and the rationale for whether they used, or did not use, the guideline. Of eighteen participants, sixteen returned questionnaires.

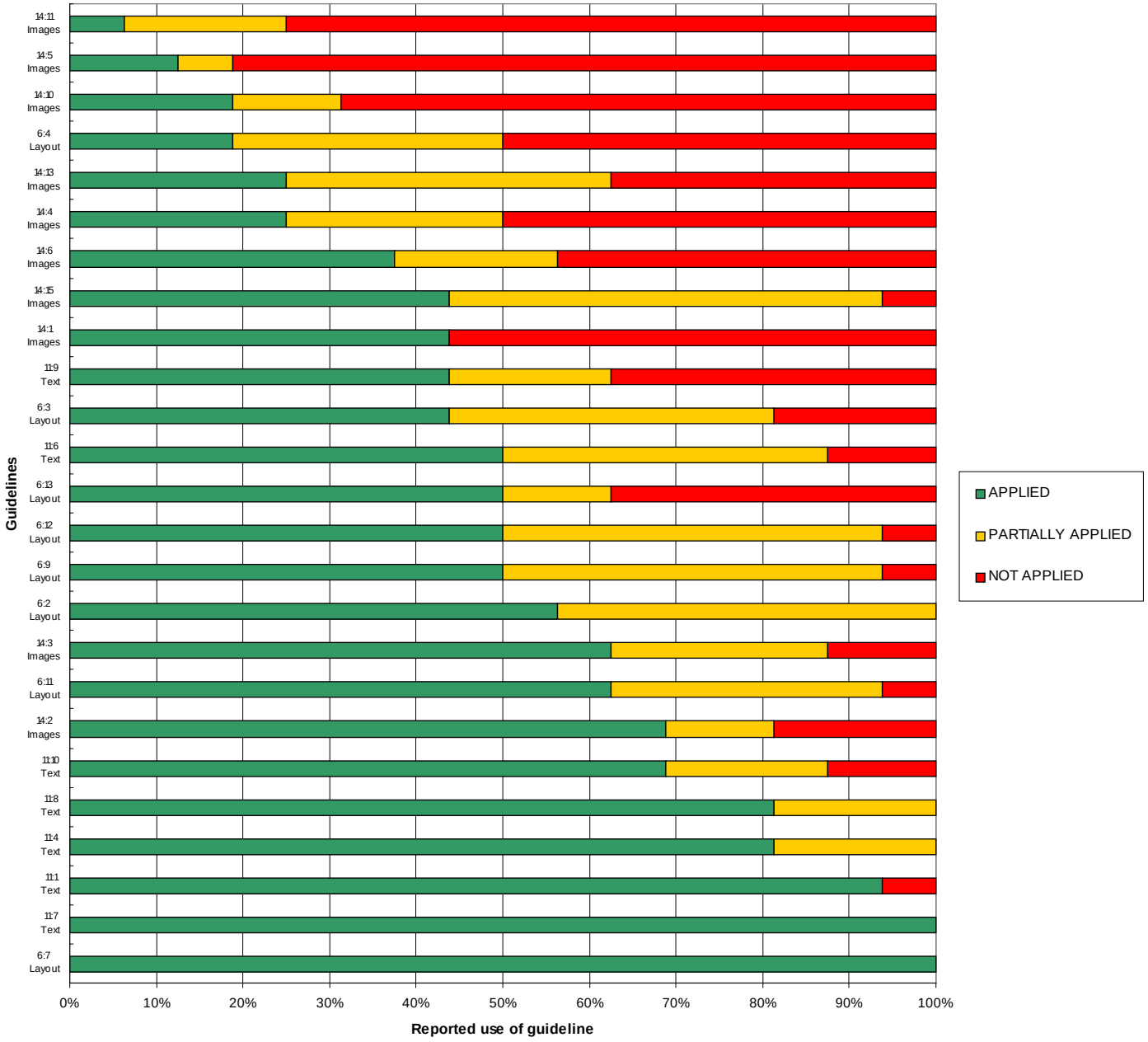
Table 1 lists the five guidelines that at least twelve of the 16 responding participants either anticipated using or had already implemented in their projects. How the guideline was interpreted and implemented is not known, although it is noteworthy that four of the guidelines are drawn from the Text category. The limitations of working within a Wiki design environment may have been a factor, since text changes are, for the most part, possible whereas more complex design choices are not.

Table 1: Top Five Guidelines used by participants

Guideline #	Category	Guideline	% of participants who used guideline
6.7	Layout	Visually align page elements, either vertically or horizontally.	100%
11.7	Text	Use a familiar font to achieve the best possible reading speed.	100%
11.1	Text	When users are expected to rapidly read and understand prose text, use black text on a plain, high-contrast, non-patterned background.	94%
11.4	Text	Ensure visual consistency of Web site elements within and between Web pages.	81%
11.8	Text	Use at least a 12-point font (e.g., typeface) on all Web pages.	81%

Figure 1 lists all of the guidelines (by number and category) and illustrates the distribution of APPLIED, PARTIALLY APPLIED and DID NOT APPLY in the response to the question of whether the particular guideline was applied to the project design. Guideline 14.11 was considered the least applicable to the designs, whereas 6.7 (Layout) and 11.7 (Text) received unanimous votes for applicability. While the limitations of Wiki was a factor in the applicability of some guidelines, the particular context used in the project was the dominant reason certain guidelines were not used. For example, projects which focused on textual information and did not use any illustrations would make irrelevent guidelines addressing the graphical representation of information. The twenty five guidelines cover a broad range of design problems, some of which were not a factor in the Wiki design project.

Figure 1: Applicability of guideline to Wiki design project



Eighty percent of the reasons given in the questionnaire for why a guideline was not applied were related to the non-applicability of the guideline. In particular, guidelines which address the use of images (14.5) or graphic representations of data (14.10, 14.11) were not applicable since the design projects did not incorporate visual data. The design limitations of a Wiki page accounted for 14% of

the reasons why guidelines were not applied. . It is worth noting that there was some overlap in how the results were reported. For example, of the twelve participants who did not apply guideline 14.5, seven felt that Wiki limitations were the reasons, while five participants felt it was because of the context.

Table 2: Reasons guidelines were not applied to the design

Reason given why guideline was not applied	Percentage of respondents
Focus of guideline was not applicable (context dependent)	80%
Limitation of the Wiki site	14%
Misunderstanding of guideline	2%
Did not agree with guideline	2%
Other	3%

There were some misunderstandings of the guidelines (where a participant checked NOT APPLIED, but in the explanation indicated that they had used the guideline. It appears that the guideline was misunderstood and the participant should have indicated APPLIED). At least two NOT APPLIED results were reported because a particular participant had not reached that point in design when the guideline might be considered. This latter category is captured as Other in Table 2. Two participants indicated a single guideline each that they did not agree with. This low reporting of guideline conflict in the questionnaire ran counter to the data collected in the focus group sessions, where a larger number of guidelines were identified as being either confusing or of limited value. The reasons for the discrepancy are unclear (and perhaps have to do with the nature of a focus group, where peers influence discussion and potentially bias the data (Sim, 1998; Lunt & Livingstone, 1996).

3.2 Focus Groups

The focus group sessions each included three student working groups (of three), for a total of nine participants. The groups were kept together in the hope that the dynamic between members would shed light on the interpretation of guidelines, a factor discussed in previous studies (Ivory &

Megraw, 2005, Cherry, Muter & Szigeti, 2006). Each session started with brief introductions and initial thoughts on the guideline set, and discussion immediately turned to the limitations of Wiki and the inability to apply a number of the guidelines. In two cases, the Wiki design had not progressed to a point where a particular guideline could be applied, although the participants completed the questionnaire and considered the guidelines as though they would be applied in the near future.

While the participants were not familiar with the source of the twenty five guidelines, they did express familiarity with the guideline contents. The term used most often to describe them was "common sense." As one participant indicated, "we've had exposure to a lot of these guidelines before so for us it's second nature", although the exposure was not to this specific set, but rather to similar ideas found in other literature. Participants expressed that this familiarity came from both design experience, either directly or indirectly through friends who have designed web sites, or from the classroom, "these ideas have been hammered into our heads for the past two years so for a lot of it we don't need the guidelines to know what to talk about. If you hadn't done design it wouldn't be so obvious." A number of participants had studied human factors issues and brought that knowledge to bear on this particular class assignment. They made reference to specific classes and instructors who had first presented some of the ideas captured in the guideline set used for the present study. Few of the participants had any design experience, but all felt comfortable assessing design usability due to previous course work. They considered themselves "experts", in contrast to users with less experience. Differences in usability experience was mentioned throughout the focus group discussions. An ideal user of the guideline set, in the words of one participant, were "people who design websites [but] aren't exposed to a fraction of the human factors design experience that we have been." They might not be termed "novice" users, but rather people, "without user interface design training and the [related] history and background." Other participants used different

expressions to describe the idea user. The common definition, regardless of the descriptive terms used, was that they lacked usability or human factors design education. The participants in the focus group had this knowledge and argued that the guideline set was of less value to them as they were to those without the same education. "I definitely think it's useful for training purposes and for anybody lacking experience designing," said one participant. This belief was echoed by others in the focus groups.

The consideration of the guidelines as "common sense" does not imply that they were redundant or without value. They reinforced existing knowledge and a number of participants suggested that the guidelines could be used as a check list against which to measure prototypes. Various comments spoke to the use of guidelines not at the start of a design exercise, but rather once a prototype had been developed, "I think I would design it first the way I think it should be and how I think other people would know how to use it, then I would look at the guidelines and see which guidelines could make this improve or better." Guidelines which were not implemented at the prototype stage could then be explored, which – as one participant suggested – would be particularly helpful in cases where a new design feature was being considered. In both focus groups participants said that the consideration of guidelines would occur only after the design was almost complete in order to retroactively find support the design choices made:

> The only reason that I would actually use these guidelines in the future, if I wasn't in school, is to justify anything I did design. I would design it first, and if I had to justify certain aspects of whatever I make, I would go to the guideline and this is what it says and…
>
> (laughter) … because someone else said so…

> … that's how I would honestly use the guideline.
>
> You had to cover your butt for any reason, the guidelines would be one way.
>
> It's because we kind of know the guidelines, so we would just use that to backup.

One participant summed up her potential use of the guidelines and their role in the design process by saying, "the only reason would be if I were to be held responsible for some decisions being made and then if it didn't work out I could go back and say, *well I did this because of this really important* … That's what research is for, to reduce blame on yourself." While this attitude might suggest a dismissive attitude regarding the guideline set, in fact the participants viewed the guidelines as an authoritative and useful document. They trusted the guidelines for two reasons; because the guidelines were perceived to have come "from a reliable source", and – most importantly for the participants – the guidelines supported existing knowledge or opinion,

> … everyone has an idea of what a good web site is. Most of us have a few web sites in our mind that we consider well designed and when we do look at these guidelines we think back and reflect on these web pages and think, *oh yeah, this does occur pretty often on the good or better web pages*. That way we can decide whether or not a guideline is good.

In addition, none of the participants wanted to return the sets following the study, arguing for the value of the tool and wanting to keep them even though the study protocol had initially asked them to return the documents.

When asked about the sources for the guidelines, none of the participants looked in detail at the research. An assumption was made that the research validated the guidelines, and that if it met the criteria of familiarity, then there would be no need to further explore the support. As one participant noted, "… if the guidelines seem reasonable to you, you wouldn't think it's required or necessary to look into the research behind it. Because if you look at it and it seems reasonable, it's already done its job. Looking into further research is just not worth the time."

If the supporting research was not a concern for the participants, the ambiguity of some of the guidelines was a matter worthy of discussion. Concerns with the clarity revolved around the specificity versus the universality of the guidelines. The manner in which they were written left the participants unsure of how they should be applied. For example, Guideline 6.11 came under considerable criticism, in that "limit the amount of white space on pages that are used for scanning and searching" was too subjective. As one participant noted, "I think to everyone that could be different. You can't quantify that number, so it would useful if you said something useful *if out this percentage of the page is content, this percentage should be white space* or something like that." The other group went even further and did not understand the guideline itself. "What adverse effect does the white space have on scanning or searching?" they asked. During the discussion, they were clearly unconvinced that this was a guideline worth considering:

> … if you're searching for stuff, wouldn't you want less stuff on the page and more white space so you can discern what you are looking for easier?

> … Yeah, they say limit, which implies that it's bad for searching. But it's the opposite. You'd want more white space.

> … They say that the more you have on the page the less scrolling that you have to do. But there would be a lot of visual clutter you have to search through. It would require more cognitive resources to find what you're looking for than just scrolling.

However, in the questionnaires, 63% of the participants said that they would apply Guideline 6.11to their design and only one respondent said that they would not. This clear disconnect between their concern with the guideline and their reported use of the guideline is curious. They were not asked during the focus group session to explain this possible inconsistency, except that the idea behind the guideline (of limiting white space and reducing clutter) resonated with the participants, while the specific guideline left them wanting more detail and advice. This speaks to a difference between considering the guideline in isolation, and trying to apply the guideline to solve a design problem.

The actual physical format of the guidelines was not considered an issue. The manner in which the guidelines were laid out in the printed format was praised. Some participants indicated a preference for an online version over a print version was suggested by some participants, although others were fine with the print version. Ideally the guidelines should be available in both formats (which they are) to adapt to the particular preferences of guideline users.

The most important factor in whether a guideline or even a set would be applied was context. The usefulness of a guideline was based on the contextual requirements of the design – if the design was text heavy and required the user to read quickly, the appropriate guideline would

address this particular design challenge. If reading speed was not an issue, related guidelines would obviously have diminished value.

If the guidelines only reinforced existing assumptions on what constituted good design, how did the participants judge good design? Participants suggested that while classroom experience helped them assess the validity of a guideline, the best examples of good design were found in existing popular sites. If a highly visited site exhibited certain design characteristics, those characteristics would constitute an example of good design. In fact the first step many of the participants took in their project work was to visit other sites to see how they were designed. They did not seek guidelines or specific design advice, but rather would look at sites with similar content in order to understand how design problems were addressed, "I personally looked at big name companies because I know they have money and therefore paid professional people to do it. So clearly they're the professionals and they know how to do it and I followed the big name websites." While it was discouraging to hear human factors engineering students defering to the notion that popularity was an indicator of good design, it is unclear based on this single study how representative such deference is. There is no clear consensus on the role that guidelines play in design, or even on the attitudes towards the use of guidelines. At the end of one focus group session, a participant perhaps best summed up the tone of the discussions,

> I think overall the guidelines in question and guidelines overall are useful. They should be part of the design process, but not necessarily exhaustive of your entire process. They are avenues you should also explore and maybe the guidelines are a starting point, but there are other avenues you should explore when designing like checking out other websites, doing a literature review, or doing usability testing iteratively.

4.0 Discussion

The present study found that guidelines played a marginal role in the design process. While the student participants had been previously exposed to guideline sets, they did not explicitly use them in the design process. Instead, they internalized what they had been taught and used examples from other sites to address design problems. There is reason to assume that a guideline set such as the one used in the study will also be internalized to some degree by the guideline users, and the desire of the participants to keep the guidelines following the study speaks to their recognition that the research conveyed in the form of guidelines was useful.

While there is no reason to doubt the expertise of the particpants in human factors, the use of guidelines as a means to support design solutions after they have been applied, if true, was unexpected. Both the questionnaire and the focus group discussions brought to light the need for tools that designers could use. What remains unclear, is when these tools would be used (prior to design, at some point in the middle, or upon conclusion) and how specific they should be to the design problem. It can be argued that the guidelines should be consulted prior to beginning design, but that is not when the participants in this study would use them. The question of the level of specificity of a guideline is very difficult to answer. Can guidelines be written representing broad research findings but include specific examples? Arguably that would not be the best use of the guidelines, since they are not rules. While some guidelines (such as the guidelines regarding the use of white space on a page) are clearly too broad, it did lead to a discussion that might result in better design, simply because the issue was raised. The amount of white space might be unclear and be context dependant, but the participants will now have the guideline in mind and may design accordingly.

While the present study has some shortcomings both in terms of the subject group and the tasks involved, collecting information regarding how guidelines are perceived and potentially used is important to understanding the value of guidelines as a means to convey knowledge. If guidelines are not used by designers, or if they are presented in a manner that makes usage difficult, then the development of guidelines should be re-evaluated. This study did not involve designers, but rather students of usability who were tasked to design web pages in a limited design environment. The use of a Wiki site may have suited the purposes of the class, but its limitations affected the data. But given these factors, the data collected was still useful partially because there is so little data in this area. This represents a start – a stake in the ground – as well as a method in which designers can be asked about their use of guidelines.

This approach should be used in future studies. Providing designers with a set of guidelines, asking them to design with an awareness of these guidelines, and then discussing how they used the guidelines would result in a better understanding of the guideline as a format for conveying design knowledge. Would designers make better use of guidelines if they were not limited by a Wiki environment? Would they also rely on expert opinion to design, and only use the guidelines as a secondary check list? The participants of the present study used a combination of experiential knowledge and the searching out of similar sites (examples) in order to design. It is worth considering whether that approach represents the norm in the design of interfaces, where guidelines may not currently have a role to play.

References

Beier, B., & Vaughan, M.W. (2003). *The Bull's Eye: A Framework for Web Application User Interface Design Guidelines.* Paper presented at the CHI 2003, Ft. Lauderdale, Florida.

Cherry, J. M. (1998). Bibliographic Displays in OPACs and Web Catalogs: How Well Do They comply with Display Guidelines? *Information Technology and Libraries, 17*(3), 124-137.

Cherry, J. M., Muter, P., & Szigeti, S. (2006). Bibliographic Displays in Web Catalogs: Does Conformity to Design Guidelines Correlate with User Performance? *Information Technology and Libraries, 25*(3), 154-162.

Cohen, A., Crow, D., Dilli, I., Gorny, P., Hoffman, H.J., Iannella, R., Ogawa, K., Reiterer, K., Ueno, K., & Vanderdonckt, J. (1994). Tools for Working with Guidelines. *SIGCHI Bulletin, 27*(2), 30-32.

D'Angelo, J., & Twining, J. (2000). Comprehension by Clicks: D'Angelo Standards for Web Page Design, and Time, Comprehension, and Preference. *Information Technology and Libraries, 19*(3), 125-135.

Henninger, S., Lu, C., & Faith, C. (1997). *Using Organizational Learning Techniques to Develop Context-Specific Usability Guidelines.* Paper presented at the Symposium on Designing Interactive Systems, Amsterdam, The Netherlands.

Henninger, S., Haynes, K., & Reith, M., W. (1995, August). *A Framework for Developing Experience- Based Usability Guidelines.* Paper presented at the Symposium on Designing Interactive Systems, Ann Arbor, Michigan.

Iannella, R. (1995). HyperSAM: A Management Tool for Large User Interface Guideline Sets. *SIGCHI Bulletin, 27*(2), 42-45.

Ivory, M., & Megraw, R. (2005). Evolution of Web Site Design Patterns. *ACM Transactions on Information Systems, 23*(4), 463-497.

Ivory, M., Sinha, R., & Hearst, M. (2001, March 1 - April 4). *Empirically Validated Web Page Design Metrics.* Paper presented at the SIGCHI 01, Seattle, Washington.

Lunt, P. & Livingstone, S. (1996). Rethinking the focus group in media and communications research. *Journal of Communication*, 46 (2). pp. 79-98.

Ray, S. D. (2002). *Web Guidelines & Usability.* Unpublished PhD, Indiana University.

Sim, J. (1998) Collecting and analysing qualitative data: issues raised by the focus group. *Journal of Advanced Nursing* 28 (2), 345–352.

Smith, S. L., Mosier, J.N. (1986). *Guidelines for Designing User Interface Software*. Bedford,

Massachusetts: Mitre.

Sutcliffe, A. G., Kurniawan, S., & Shin, J-E. (2005). A method and advisor tool for multimedia user interface design. *International Journal of Human-Computer Interaction, 64*, 375-392.

Tetzlaff, L., & Schwartz, D.R. (1991). *The Use of Guidelines in Interface Design.* Paper presented at the Proceeding of the Conference on Human Factors in Computing Systems (CHI '91), New York, New York.

Vanderdonckt, J., Mariage, C., Scapin, D., Leulier, C., Bastien, C., Farenc, C., Palanque, P., & Bastide, R. (2000, June 19). *A Framework for Organizing Web Usability Guidelines.* Paper presented at the 6th COnference on Human Factors & the Web, Austin, Texas.

Wiki (2002) *What is Wiki.* What IsWiki. Retrieved July 18, 2007 from http://www.wiki.org/wiki.cgi