ONLINE LEARNING ENVIRONMENT DESIGN: A HEURISTIC EVALUATION

Emily Hildebrand[1], Jennifer M. Bekki[1], Bianca L. Bernstein[2], Caroline J. Harrison[3]
[1]College of Technology and Innovation, [2]School of Letters and Science, [3]School of Sustainability
Arizona State University

Abstract

The purpose of this paper is to demonstrate the feasibility of using a heuristic evaluation to evaluate the quality of and suggest improvements to an online learning environment. The website evaluated in this study is a resilience training website targeting doctoral students in science, technology, engineering, and math (STEM). Findings from the heuristics evaluation are shown to provide a baseline of potential human factor issues related to user experience with the website and to suggest some preliminary design recommendations. Implications of utilizing a heuristic evaluation to improve the user-centered design of an online learning environment are further discussed.

Introduction

The use of educational technology and online learning environments (OLEs) are increasing at a rapid rate. In fact, it is estimated that by 2025 there will be between 30-80 million online students [1]. Not only is student enrollment going up, but the types of online resources available to students are expanding. For example, there are online learning systems that fall into the category of course management systems (e.g., Blackboard and Moodle), massive open online courses (MOOCS), as well as an increasing number of informal online learning environments that provide support to more structured, class-based learning. The growing use of and variety of formats for online learning both demonstrate its value and underscore the fact that its use is increasingly necessary in order to meet the demands of today's learners.

Given their widespread use, the design principles and usability evaluation metrics for OLEs are of great interest. Recent research has suggested that the design of an OLE can influence the student learning experience [2]. Further, poor design and usability have been found to lead to high rates of attrition [3, 4]. These findings are not surprising given that OLEs are highly variable in terms of accessibility, formatting, and levels of user control. While much of the literature is calling for the development of design guidance [2,5,6,7,8], there is still much to learn about best practices when it comes to designing for optimal usability in OLEs, and it is not yet known what works best [9,10]. With a field that is fast paced and continually changing to meet educational needs, the research to validate best practices has not caught up, creating the fear that "today's best practices may soon be embedded in antiquated technology" [9].

In human factors engineering, practitioners often rely on the tenets of user centered design to optimize the efficiency of users' tasks and performance. However, the literature suggests that there may be a distinct difference when designing for basic user tasks and tasks to facilitate learning [10]. It is not only important to recognize that there is a difference, but to understand the characteristics, expectations, and needs of this unique "learner" population. Quintana, *et al.* (2003) noted that traditional user centered design is focused on helping people complete specific tasks that they typically already know, whereas learner centered design is focused on helping people develop an understanding of novel knowledge. Thus, integrating theories from pedagogical design with human factors design principles will be critical for achieving the greatest learning outcomes from an OLE [2, 10].

While the users and developers of OLEs lack agreement about what learner centered design involves and/or how it should be undertaken [1], there are many usability evaluation methods in human factors engineering that could provide a useful starting point. Traditionally, these methods are used as an efficient way to understand where design may be creating difficulty for end users and impacting performance. Once trouble areas are established and evaluated across OLE disciplines, general solutions can be developed. One methodology that has already been suggested as appropriate for evaluating the quality of an OLE is the heuristic evaluation [2, 4, 8, 11]. While there are numerous usability methods, heuristic evaluations are popular for addressing user interfaces because they are cost-effective, efficient, and easy to apply [4, 8, 11, 12].

Developed by Nielsen [11, 12], a heuristic evaluation involves expert raters applying a usability checklist to a user interface in order to identify potential usability problems that may result in users not finding the information they need or not being able to execute their desired task. A traditional heuristic evaluation primarily involves comparing the user interface with a checklist made up of ten general design criteria [11, 12], or "rules-of-thumb," that should be followed when designing a product or process in order to optimize the user experience.

Case Study

The primary goal of this study was to demonstrate the feasibility of using a heuristic evaluation to gauge the quality of and suggest improvements to a psychological education website (http://careerwise.asu.edu) [13]. The *Career*WISE resource is an OLE designed to provide resilience training to women doctoral students in the fields of Science, Technology, Engineering, and Mathematics (STEM). For the purposes of this research, *Career*WISE serves as a prototype, providing a vehicle to demonstrate the proof-of-concept for how different approaches to user-centered design could be applied to an OLE. Findings from the case-study heuristic evaluation are shown to provide a baseline of potential human factors issues related to user interaction with the *Career*WISE website and to suggest some preliminary design recommendations.

Methods

Participants

Five expert raters conducted a heuristic evaluation of the *Career*WISE website. The raters were all female graduate students with a background in human factors and previous experience with the foundations of the heuristic evaluation methodology.

Materials and Procedure

The procedure for this study followed the guidelines and heuristics (shown in Table 1) established by Nielsen [11,12] for evaluating user interfaces. When completing a heuristic evaluation, it's important to have the raters carry out tasks that are most common for typical users to do when interacting with the website. This ensures that the rater gets a comprehensive look at the user interface and its functionality. In this study, each rater was asked to review the *Career*WISE website by completing the following standard user tasks:

1. Go to careerwise.asu.edu, create a free account, and login.

2. Click through all steps of the problem solving model.

3. View a HerStory video from someone in the Physics field and read the biographical information for the interviewee.

4. Find the "Make good decisions" skill page. From that page, watch a HerStory video clip that is related to the common concern "Advisor Issues".

5. Search for information on "starting a family in grad school"

HEURISTIC EVALUATION WORKSHEET				SEVERITY						
EVALUATOR NAME:				I don't agree that this is a usability problem.	Cosmetic Problem Only	Minor Usability Problem - Low Priority	Major Usability Problem - Major Priority	Usability Catastrophe - Fix prior to product release	Potential Solution(s)	
DATE:										
SYSTEM: CareerWISE										
TASK #	HEURISTIC #	LOCATION / PAGE TITLE	DESCRIPTION	1	2	3	4	5		
1	3	Homepage	Users cannot see past the first page without logging in or creating an account.				X		Don't require users to create an account, or make it optional.	EXAMPLE

Figure 1. Raters used the above worksheet to record heuristic violation, severity ratings, and design recommendations.

6. Spend 30 minutes freely searching through the site.

Using the standardized worksheet shown in Figure 1, the five raters independently reviewed the *Career*WISE website, completing the five common tasks in order to identify violations of the heuristics shown in Table 1. Each rater completed a separate worksheet, listing the task they were completing when they noticed the violation, the heuristic from Table 1 that was violated, the location of the violation within the website, and a description of the problem.

Following the initial evaluation, the worksheets from each of the reviewers were compiled into a single master list of violations by the experimenter, and redundancies were edited out. Any usability problem that was found by multiple reviewers to violate the same heuristic was considered a redundancy. Each of the original raters then reviewed the master list independently and assigned severity ratings to the violated heuristics. Wherever possible, they also provided a redesign recommendation. The severity ratings were based on the magnitude and criticality of the usability problems that could occur as a consequence of the heuristic violation and were made based on the following scale:

1 = I don't agree that is a usability problem
2 = Cosmetic problem only
3 = Minor usability problem – low priority
4 = Major usability problem – high priority
5 = Usability catastrophe – Fix prior to product release

The results for this study are presented in the following section in several distinct ways. First, the frequency of the violations found for the *Career*WISE website is summarized according to the ten categories of usability heuristics shown in Table 1. Second, the severity ratings across all violations are summarized according to the frequency with which they were assigned (e.g., a violation of severity level 5 was identified x number of times). Third, a summary of the data according to the frequency, severity, and location of all heuristic violations is presented. Finally, we provide specific examples from the *Career*WISE website for three heuristics (from Table 1) that were found to be violated the most frequently.

Results

In total, there were 50 unique usability problems identified within the *Career*WISE website. Since each usability problem could violate more than one heuristic, there were 91 actual heuristic violations. Figure 2 displays the frequency of violations according to heuristic category (Note: heuristic abbreviations are defined in Table 1).

Three heuristics accounted for 66% of the total violations. Match, which evaluates the corresponddence between the website and the real

Table 1: Nielsen heuristics [12].

#	Heuristic	Abbreviation	Notes
1	Visibility of system status	Visibility	· The website keeps the user informed about what is going on through constructive, appropriate and timely feedback.
2	Match between the system and the real world	Match	· Language usage, such as terms, phrases, symbols, and concepts, is similar to that used by the users in their day-to-day environment. · Information is arranged in a natural and logical order.
3	User control and freedom	Control	· Users control the system. · Users can exit the system at any time even when they have made mistakes. · There are facilities for Undo and Redo.
4	Consistency and adherence to standards	Consistency	· Concepts, words, symbols, situations, or actions refer to the same thing. · Common platform standards are followed.
5	Error prevention, specifically prevention usability-related errors	Error	· The system is designed in such a way that the users cannot easily make serious usability errors. · When a user makes an error, the application gives an appropriate error message.
6	Recognition rather than recall	Recognition	· Objects to be manipulated, options for selection, and actions to be taken, are visible. · The user does not need to recall information from one part of a dialogue to another. · Instructions on how to use the system are visible or easily retrievable whenever appropriate.
7	Flexibility and efficiency of use	Flexibility	· The site caters to different levels of users, from novice to experts. · Shortcuts or accelerators, unseen by the novice users, are provided to speed up interaction and task completion by frequent users.
8	Aesthetic and minimalism in design	Aesthetics	· Site dialogues do not contain irrelevant or rarely needed information, which could distract users as they perform tasks. · Displays are simple and multiple page displays are minimized.
9	Recognition, diagnosis, and recovery from errors	Recovery	· Error messages are expressed in plain language. · Error messages indicate precisely what the problem is and give quick, simple, constructive, specific instructions for recovery.
10	Help and documentation	Help	· The site has a help facility and other documentation to support the users' needs. · The information in these documents is easy to search, focused on the user's task, and lists concrete steps to be carried out to accomplish a task.

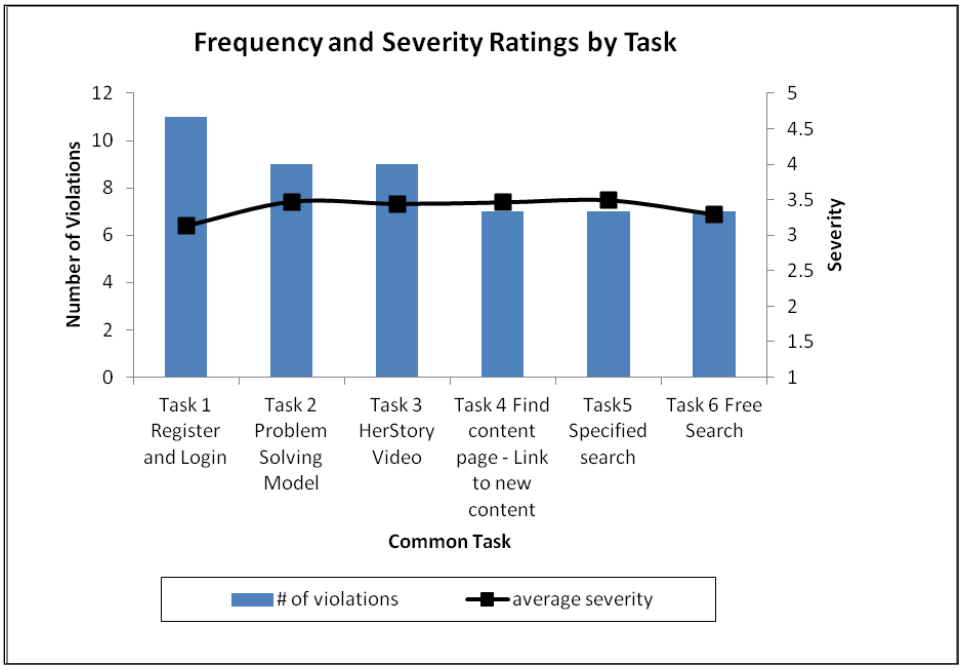

Figure 2. Frequency and Average Severity ratings by Common Task.

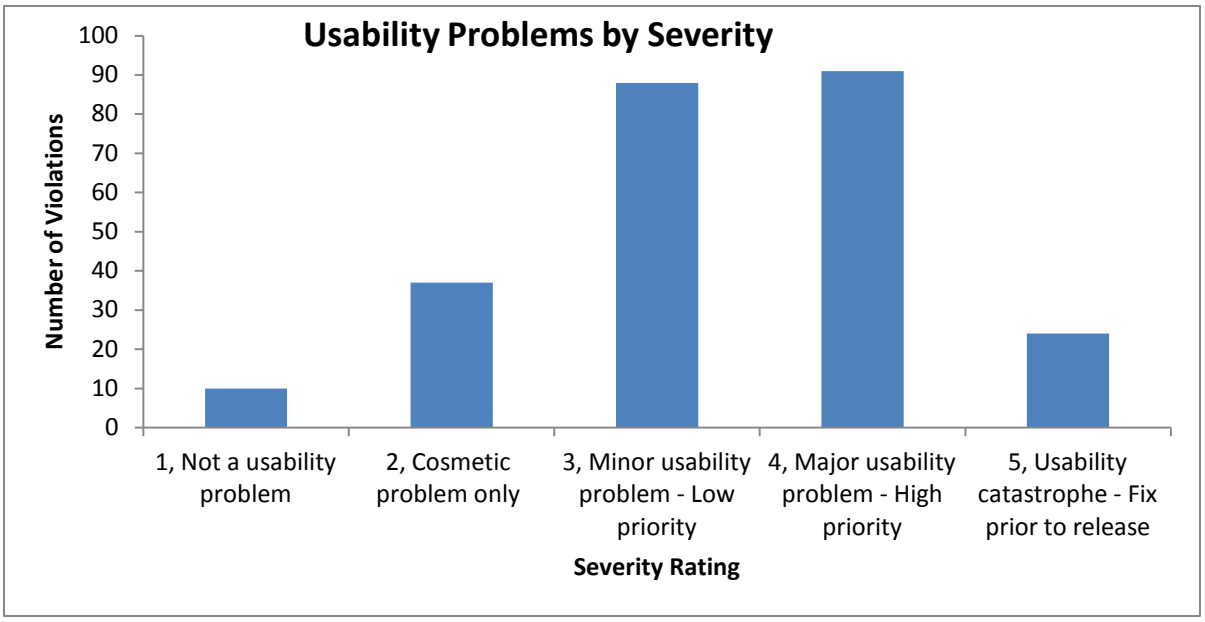

Figure 3. Frequency of severity ratings applied to usability problems.

world, was the most violated heuristic (24 times), with Consistency and Aesthetics following closely behind (20 and 16 respectively). The Recovery heuristic, which assesses the appropriateness of error messages, was found to have no violation within the *Career*WISE website.

Figure 3 shows the frequency of the severity ratings assigned to the usability problems. Of the 50 total usability problems identified, the raters found 81% to be between a minor usability problem and a catastrophic problem. For example, one usability problem rated as severe was the lack of consistent navigation making it difficult for the user to know what page they had last come from. Conversely, 19%

of problems were found to be either not a usability problem or only a cosmetic issue. An example of a cosmetic issue that the raters identified was inconsistent header alignment on some of the pages. Overall, the severity findings are notable because, typically, when a large percentage of the usability problems are rated as moderate to severe, it implies that the user's experience and, perhaps more importantly in the case of an OLE, the user's learning from the site, may be suffering.

Finally, to give an overall view of the results, Figure 4 combines the frequency, average severity, and respective common task where the heuristic violation was found within the *Career*WISE website. There was little variance in the number of violations identified by raters across the common tasks. Similarly, severity was not focused in one area over another, implying that usability issues may be affecting the user experience across the website instead of being focused in a single task area. The results shown in Figure 4 make it apparent that usability problems were not inherent in a single task, but were found throughout the website. To begin addressing the findings, it is often the best strategy to remedy those issues with the highest severity ratings and those issues which continually violate multiple heuristics at a time. The following section will highlight some examples of the most severe heuristic violations and discuss possible strategies to improve the design.

Example Heuristic Violations

As mentioned previously, the three most frequently violated heuristics were "Match", "Consistency", and "Aesthetics". Following are examples of a specific violation for each of the three heuristics.

Match

A violation of the heuristic "Match between the system and the real world" signifies that language, terms, and phrases do not "match" the terms employed by users on a daily basis or that the information is arranged in an unnatural or illogical order. During the heuristic evaluation, multiple raters noted that at times the *Career*WISE website used terms that were unfamiliar. For example, the term "Briefs", which is used within the *Career*WISE site to represent a page with a short write-up related to

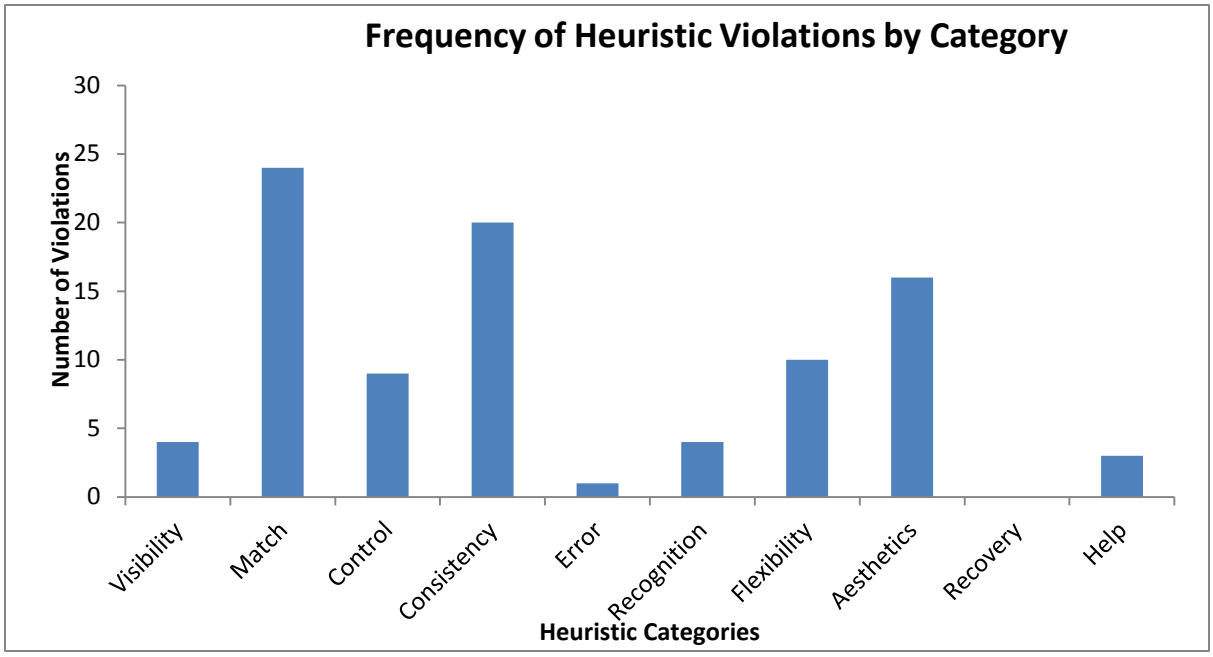

Figure 4. Frequencies of the heuristic violations by category averaged across raters.

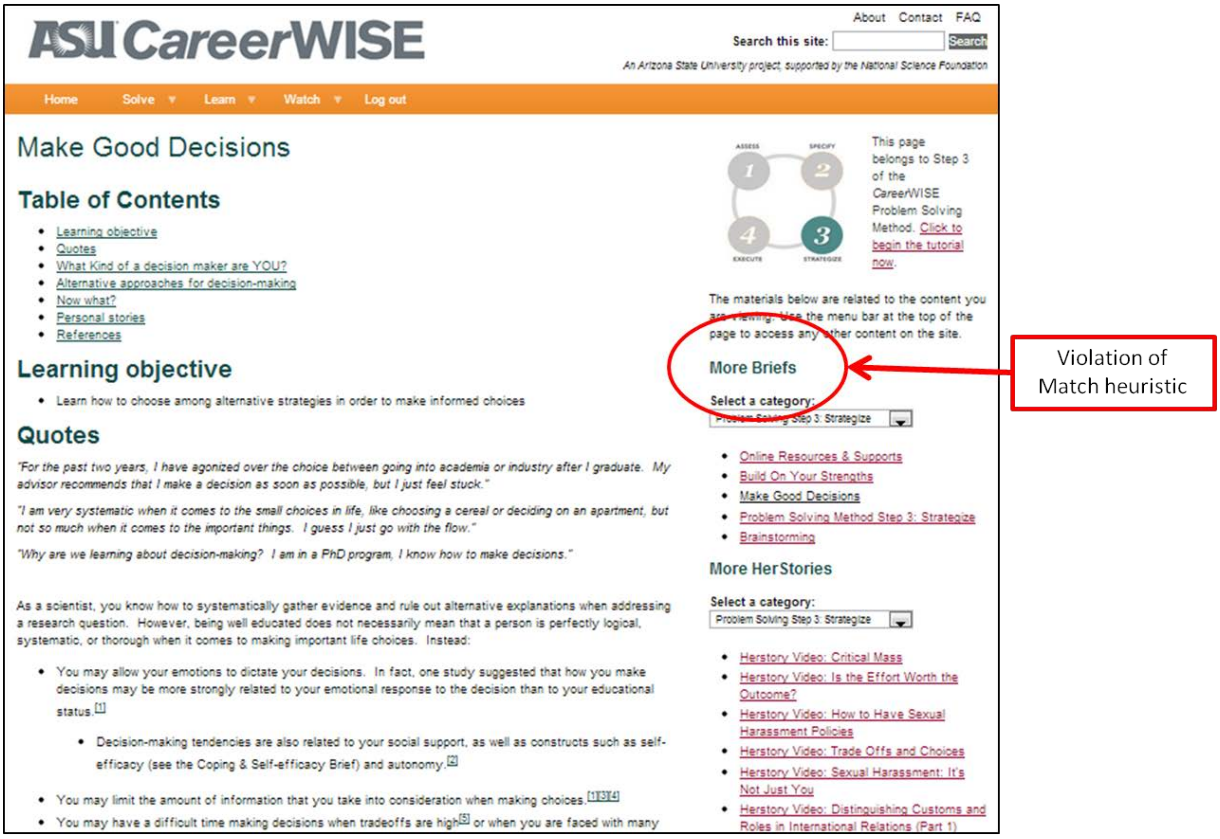

Figure 5. A Screenshot from the CareerWISE website shows a violation of the Match Heuristic.

a particular topic, was found to be confusing to raters. Figure 5 provides a screen shot highlighting such a violation. A potential solution to this usability problem could be to work with a focus group comprised of end users to identify more appropriate and familiar language that would better represent that type of content on the website.

Consistency

A violation of the heuristic "Consistency and adherence to standards" most typically occurs when the same concepts, words, symbols or actions do not refer to the same thing. An example of this that occurred during this heuristic evaluation was that the *Career*WISE homepage has two links within the same textbox that are labeled the same, but when clicked lead to different locations within the site. This is highlighted by the two links title "More" in the screen shot provided with Figure 6. To improve upon this usability problem it is important to ensure that the two links, which lead to different content, are easily distinguishable from one another. This could be accomplished by changing the text, color, and/or location of the link.

Aesthetics

A violation of the heuristic "aesthetic and minimalism in design" occurs when site dialogue contains irrelevant information that is distracting to users, when displays are overly complex, and/or when page displays are not minimized. For example, during this heuristic evaluation, there were multiple times when menus and search results required an excessive amount of pages for the user to search through in order to access information. Figure 6 shows a menu of related videos that spans 6 pages. To remedy this usability problem, it is best to consolidate all links for the same "menu" on one page. Human factors design principles can offer many strategies to accomplish this, for example using nested menus.

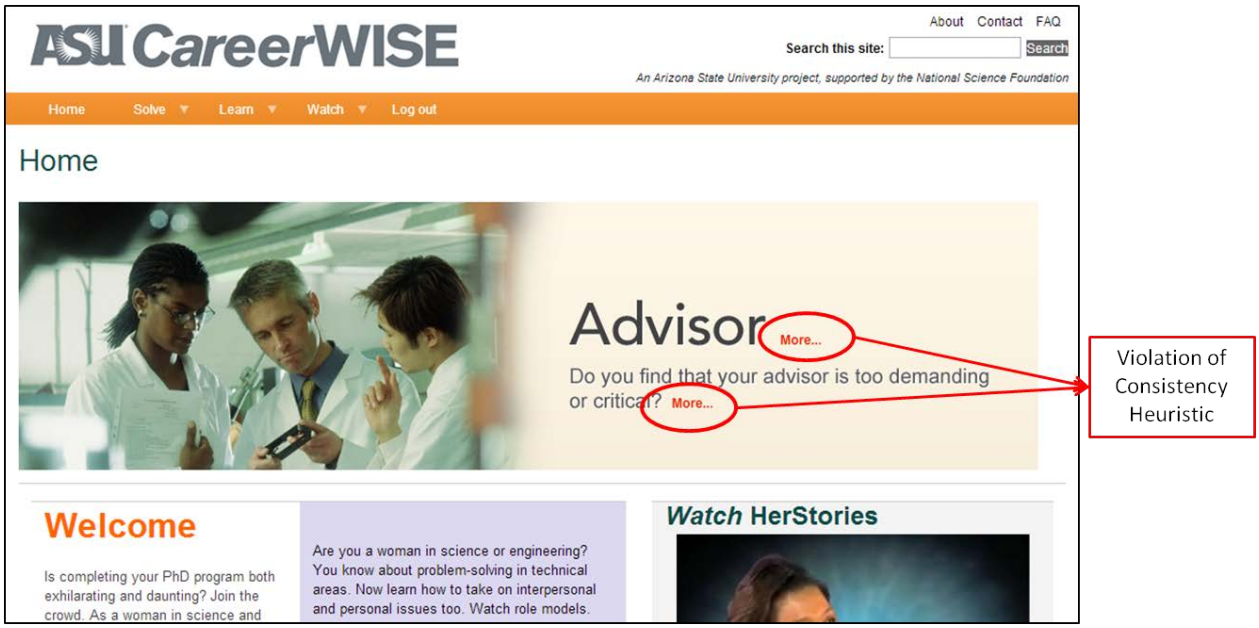

Figure 6. A screenshot from the CareerWISE website shows a violation of the Consistency Heuristic.

Discussion and Conclusions

The present study explored the application of usability evaluation methods in an online learning environment (OLE). Specifically, we intended to demonstrate the feasibility of heuristic evaluation as a methodology for evaluating an OLE. The *Career*WISE OLE was used as a test case for this study, and a baseline of potential human factors issues related to the site were identified. With the three most violated heuristics being Match, Consistency, and Aesthetics, this evaluation indicated that *Career*WISE may have many potential issues related to a mismatch between the existing interface and user expectations for site usability. Furthermore, since none of the common tasks completed by the raters during the heuristic evaluation acquired a majority of violations, the data also suggests that the problems are more inherent within the overall structure of the website than within specific task areas. If an OLE designer lacked time and funds to do further research, the results from a heuristic evaluation could highlight potential problem areas that could then be amended using general user-centered design principles, as demonstrated through this study. However, if additional time and/or funds were available, follow-up usability studies should focus on what and where these "mismatches" are within the website.

Of note is that while the heuristic evaluation confirmed that the current user interface may have a number of potential usability problems, we still do not know which problems pose negative consequences for learning outcomes. Other research shows that OLEs that are not designed to meet student needs can result in suboptimal learning outcomes by users [2, 3, 4], so it will be important moving forward to identify frameworks and guidelines that can inform the OLE design principles that best support learning. Thus, a next important step in this research will be to perform formal usability tests and observation-based studies to determine which problems are in fact most troublesome with relation to key learning outcomes.

While findings from this study are in fact specific to the *Career*WISE website, the process followed is applicable to any OLE. However, the findings from an evaluation of another OLE could be completely different in terms of the types and severity of heuristic violations. As such, more research is needed across platforms and domains in order to continue developing a "baseline" of human factors issues present in

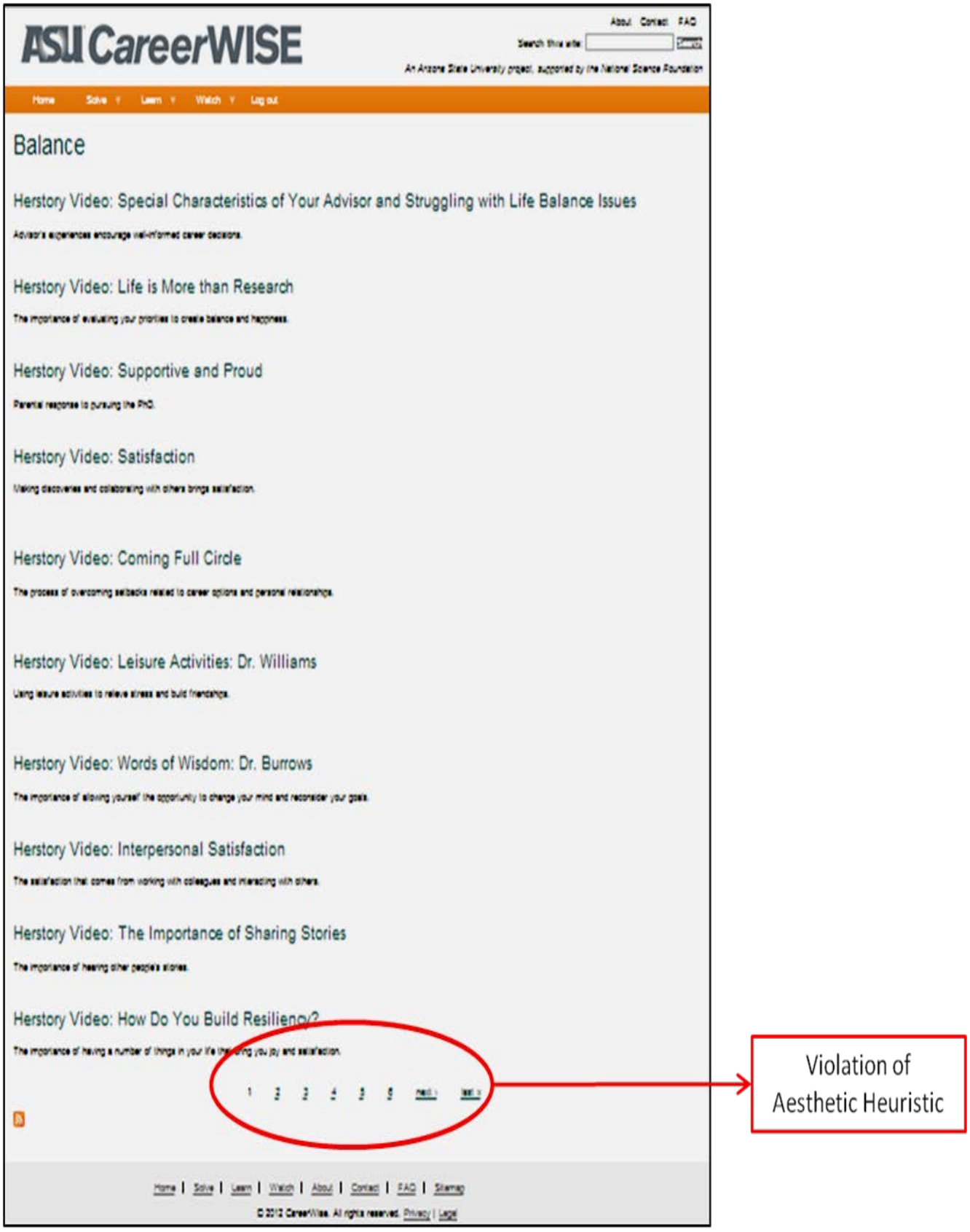

Figure 7. A screenshot from the CareerWISE website shows a violation of the Aesthetic Heuristic.

OLEs to aid website designers in improving existing and newly created OLEs.

Finally, there are many "bigger picture" questions that need to be answered in order to truly improve both the usability of and resultant learning from OLEs. First, as noted by Quintana, *et al*. (2003), there needs to be a shift from user centered design to learner centered design. In order to develop effective design principles, frameworks and guidelines, the goals, expectations, and needs for these "learners" need to be established [2, 5, 7, 10]. Furthermore, beyond establishing criteria for one type of learner or OLE, research suggests that a paradigm shift to "universal design for learning" is necessary [14] in order to meet the diverse needs of a wider variety of learners across different OLE domains.

Acknowledgments

This study was supported by the National Science Foundation (NSF) grant 0910384. Any opinions, findings, and conclusions and recommendations expressed in this report are those of the authors and do not necessarily reflect the views of the NSF.

Bibliography

1. Hosie, P., Schibeci, R., & Backhaus, A. (2005). A framework and checklists for evaluating online learning in higher education. *Assessment & Evaluation in Higher Education*, 30(5), 539-553.

2. Giering, J. A. (2012). *Use of evaluation to design quality online learning: understanding the shared experience* (Doctoral dissertation, Drexel University).

3. Minocha, S., & Sharp, H. (2004). Learner-Centered and Evaluation of Web-Based E-Learning Environments. *The 7th HCI Educators Workshop: Effective Teaching and Training in HCI.* Preston, United Kingdom.

4. Ssemugabi, S., de Villiers, M.R. (2010). Effectiveness of heuristic evaluation in usability evaluation of e-learning applications in higher education. *South African Computer Journal*, 45, 26-39.

5. Cook, D.A., & Dupras, D.M. (2004). A practical guide to developing effective web-based learning. *JGIM*, 19, 698-707.

6. Hoyle, S., Bruton, K., Peres, S.C., & Gutierrez, T. (2012). Usability of an Interactive Educational Website for Statistics. *Proceddings of the Human Factors and Ergonomics Society 56th Annual Meeting,* 1827-1831

7. Ssemugabi, S., & de Villiers, R. (2007). A comparative study of two usability evaluation methods using a web-based learning application. *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, Garden Route, Wilderness, South Africa, October.

8. Zaharias, P., & Koutsabasis, P. (2011). Heuristic evaluation of e-learning courses: a comparative analysis of two e-learning heuristic sets. *Campus-Wise Information Systems*, 29(1), 45-60.

9. Hines, P.J., Jasny, B.R., & Mervis, J. (2009). Adding a T to the Three R's. *SCIENCE*. 323, 53.Nielsen, J. (1993). *Usability Engineering*.Cambridge, MA: Academic Press Professional.

10. Quintana, C., Krajcik, J., & Soloway, E. (2003). A Framework for Understanding the Development of Educational Software. In Jacko, J.A., & Sears, A. (Eds.), *The Human-Computer Interaction Handbook*, (823-834). Mahwah, NJ: Lawrence Erlbaum Associates.

11. Nielsen, J. (1992). Finding Usability Problems through Heuristic Evaluation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 373-380.

12. Nielsen, J. (1993). *Usability Engineering*.Cambridge, MA: Academic Press Professional.

13. Bernstein, B. L. (2011). Managing barriers and building supports in science and engineering doctoral programs: Conceptual underpinnings for a new online training program for women. *Journal of Women and Minorities in Science and Engineering. 17* (1), 29-50.

14. Poore-Pariseau, C. (2010). Online Learning: Designing for All Users. *Journal of Usability Studies*, 5(4), 147-156.

Biographical Information

Emily Hildebrand is currently a PhD student in Simulation, Modeling, and Applied Cognitive Science at Arizona State University and a Research Assistant in the CareerWISE research program, an online resilience training program for female graduate students in STEM. Ms. Hildebrand's research interests include human factors, usability, and user experience.

Jennifer M. Bekki is an Assistant Professor in the Department of Engineering & Computing Systems at Arizona State University. Her research interests include educational data mining, online learning, engineering student persistence, STEM graduate students (particularly women graduate students), discrete event simulation methodology, and the modeling and analysis of manufacturing systems.

Bianca L. Bernstein is Professor of Counseling and Counseling Psychology in the School of Letters and Science at Arizona State University. Her research focus is on the application of psychological science to the career advancement of women and underrepresented minorities and the development of effective learning environments for graduate education.

Caroline J. Harrison is currently the Curriculum Developer for the School of Sustainability at Arizona State University. Her current work involves the development of innovative courses and programs serving undergraduate and graduate students. Her research background is in the area of online, multimedia instruction.