Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Biomedical Informatics 39 (2006) 412–423
Journal of Biomedical Informatics
www.elsevier.com/locate/yjbin

Heuristic evaluation of paper-based Web pages: A simplified inspection usability methodology

Mureen Allen [a,*], Leanne M. Currie [a,b], Suzanne Bakken [a,b], Vimla L. Patel [a], James J. Cimino [a]

[a] *Department of Biomedical Informatics, Columbia University, New York, NY, USA*
[b] *School of Nursing, Columbia University, New York, NY, USA*

Received 27 April 2005
Available online 15 November 2005

Abstract

Online medical information, when presented to clinicians, must be well-organized and intuitive to use, so that the clinicians can conduct their daily work efficiently and without error. It is essential to actively seek to produce good user interfaces that are acceptable to the user. This paper describes the methodology used to develop a simplified heuristic evaluation (HE) suitable for the evaluation of screen shots of Web pages, the development of an HE instrument used to conduct the evaluation, and the results of the evaluation of the aforementioned screen shots. In addition, this paper presents examples of the process of categorizing problems identified by the HE and the technological solutions identified to resolve these problems. Four usability experts reviewed 18 paper-based screen shots and made a total of 108 comments. Each expert completed the task in about an hour. We were able to implement solutions to approximately 70% of the violations. Our study found that a heuristic evaluation using paper-based screen shots of a user interface was expeditious, inexpensive, and straightforward to implement.

Keywords: Clinical information systems; Heuristic evaluation; Information needs; Internet; Medical errors; Usability assessment; Web page design

1. Introduction

A number of informatics-based approaches have been proposed as part of a systems strategy to prevent medical errors and improve patient safety. As clinicians increasingly use computer and Web-based resources, attention should be focused on ensuring that the design of the user interface itself does not predispose clinicians to making potential medical errors. Indeed, poorly designed user interfaces add cognitive demands on the users and their ability to perform tasks adequately [1], and potentially could lead to errors [2,3].

The burgeoning desire for the acquisition of online medical information suggests that attention should be paid to the development and design of user interfaces with which the clinician interacts. To facilitate the presentation of online information in an effective manner, it is essential to design and develop Websites for the health care domain that are user-friendly, and take into account the users' needs. Moreover, it has been suggested that academic medical institutions should "consider applying the usability methodology and formal usability evaluations" to assess their Websites [4]. However, formal usability inspection testing, which often requires a detailed analysis of user behavior, can be onerous to perform. Instead, many usability evaluators use the *heuristic evaluation* (HE) methodology to assess user interfaces, especially during the initial development stages.

2. Background

2.1. Heuristic evaluation

HE is a usability engineering method "for finding usability problems in a user interface design by having a small set

* Corresponding author. Fax: +1 908 925 4189.
E-mail addresses: mureen.allen@dbmi.columbia.edu, mallen_md@hotmail.com (M. Allen).

1532-0464/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2005.10.004

Table 1
Severity rating of heuristic violations

Severity rating[a]	Definition
0	I do not agree that this is a usability problem at all
1	Cosmetic problem only: need not be fixed unless extra time is available on project
2	Minor usability problem: fixing this should be given low priority
3	Major usability problem: important to fix, so should be given high priority
4	Usability catastrophe: imperative to fix this before product can be released

[a] Nielsen [9].

of evaluators examine the interface and judge its compliance with recognized usability principles (the "heuristics")" [5]. This method uses evaluators to find usability problems or violations that may have a deleterious effect on the user's ability to interact with the system. Typically, these evaluators are experts in usability principles, the domain of interest, or both (so-called "double" experts). Nielsen and Molich [6] described the HE methodology as "cheap," "intuitive," "requires no advance planning," and finally, "can be used early on in the development process." Often it is used in conjunction with other usability methodologies to evaluate user interfaces [7].

Furthermore, HE's utility lies in its ability to rapidly find more usability problems, including more of the major problems, compared to other methods of evaluation [8]. By evaluating the interface in the development phase, it is possible to identify design flaws. Finding these flaws earlier, rather than later, reduces subsequent usability errors, which may be more costly and prohibitive to rectify. Indeed, use of the HE methodology is ideal in the spiral or iterative development environment commonly found in the systems design industry.

The classically described HE method delineated by Nielsen et al. involves experts independently assessing the user interface of an existent device, an early prototype, or a paper model.[1] As the experts walk through the evaluation process, they identify those problems that will affect the user's ability to interact with the system. During a typical evaluation, all heuristic violations are aggregated and then the experts are asked to rank each violation according to its perceived severity (Table 1) [9].

Inspection usability methods, including the HE methodology, have been used in the medical domain to evaluate clinical information systems (CISs), Websites, and devices [10–13]. Kushniruk and Patel [13] extensively discuss the evaluation methodologies used in the domains of cognitive science and usability engineering. Additionally, they discuss how these methods can be applied to the evaluation of CISs. Zhang et al. investigated the usability of infusion pumps by conducting an HE of the interface. They postulated that the problems they found with the pump interface design could be potential sources of medical errors. In the paper by Graham et al., infusion pumps used in the intensive care unit were evaluated for usability impediments. They found major potentially catastrophic problems in a number of their evaluation categories and recommended that "end users must be informed that there are numerous aspects of the system where they need to be vigilant about the potential for making errors." Therefore, the evaluation of user interfaces within the context of the medical domain is of critical significance.

[1] A paper model usually refers to using paper implements that are moveable on a work surface that mimics a screen.

2.2. Infobutton project

We have recently described the Infobutton project, which addresses the issue of information needs while using the Web-based clinical information system (WebCIS) present at Columbia University Medical Center and New York Presbyterian Hospital [14–16]. In summary, the Infobutton project seeks to provide suitable online information resources to the end user, using contextual links or infobuttons that answer the user's information needs. The infobutton takes the user's current context (i.e., institution, user type, patient age, patient gender, clinical task, and concept of interest) and directs the user to online resources that provide solutions to the user's information needs. Behind the scenes, the Infobutton Manager (IM), a Web-based application, takes the contextual information presented to it by the infobutton, matches the information to a Context Table, and then generates a number of potential information needs (in the form of questions) and potential online information resource solutions (in the form of links) subsequently presented to the user.

Based on the results of our earlier study, we were able to delineate and categorize information needs events as they occurred in the clinical context while clinicians were using WebCIS. We used the knowledge we gained from the observational study to develop or identify online solutions to the information needs events we detected during the observational study. Subsequently, using the infobuttons and the EVI, we wished to incorporate a number of these solutions into the existing CIS's Web pages and, additionally, to develop new Web pages that incorporated other solutions to address identified information needs events.

However, before the deployment of these Web pages in WebCIS, we needed a technique to evaluate them, particularly to ensure that we presented the information needs solutions in a manner that was acceptable to the typical user of WebCIS. Though there are a number of usability techniques available (e.g., verbal protocol analysis [17], cognitive task analysis [18], cognitive walkthrough [19],

and formal user testing, etc.), we chose to use the HE technique—because of its positive characteristics: fast, inexpensive, and efficient—to judge our interfaces. At the time, we wished to modify existing, and introduce new, user interfaces. Therefore, we chose to design screen shots of the potential interfaces we would use for presentation to the usability experts.

This paper describes the methodology used to develop a modified HE suitable for the evaluation of screen shots of Web pages including the development of a HE instrument and thus results of the evaluation of the aforementioned screen shots. In addition, this paper presents examples of the process of categorizing problems identified by the HE and technological solutions identified for those problems. The aim of this method is to provide a more expeditious, accurate, and straightforward method to evaluate a user interface and detect deficiencies even before the device or the prototype has been developed.

3. Simplified heuristic evaluation

We adapted the heuristic evaluation methodology to apply it to the assessment of paper-based screen shots of Web pages. In this section, we describe the development of the instrument used to conduct this modified heuristic evaluation and the results.

3.1. Methods

In their work, Zhang and colleagues [11] incorporated Nielsen's 10 heuristics [20], Shneiderman's eight golden rules [21], and the results of their research to formulate a list of 14 heuristics. Three members of our group were chosen to review these heuristics (Table 2) and to decide which were the most pertinent to consider when evaluating a screen shot of a user interface. Heuristics that were agreed on by all three members were used in the subsequent evaluation process.

After the relevant heuristics were identified, 18 screen shots of WebCIS Web pages were captured. These Web pages represented observed information needs events from our previous study. Subsequently, these screen shots were modified so that they contained links to information resources accessed using infobuttons and the Infobutton Manager [22]. The screen shots were collated into a booklet, along with 14 other screen shots, showing typical Web pages that could be viewed by a WebCIS user. The evaluators were also given a narrative describing the user interacting with WebCIS as he or she accessed various components of the system—this was done to provide the evaluator with a context as they reviewed the screen shots.

An accompanying instruction manual explained the nature of the project, the Web pages that were of interest, and finally a brief description of the evaluation process and the heuristics of interest for the evaluation. The final document prepared was a rating scale for the severity of a potential heuristic violation. To expedite the process, as compared to the typical HE methodology, the evaluators assessed the severity of the problems as they identified them. As noted earlier, evaluators in HE studies usually rate the severity of all the heuristic violations including those of the other expert evaluators. By asking the evaluators to rate the heuristic violations as they found them, our study eliminated the need to re-contact the evaluators to complete the severity assessment of the aggregated heuristic violations. Based on previous work by Nielsen and Molich [6], three to five evaluators are deemed to be sufficient to detect the majority of usability problems, although this is being debated [23]. With this in mind, we choose four evaluation experts experienced in usability testing to perform the heuristic evaluation of the screen shots of our Web pages. Two of our evaluators had cognitive science backgrounds and two had clinical backgrounds.

Table 2
Definition of heuristics

Heuristic[a]	Definition
1. Consistency	The users should not have to wonder whether different words, situations or actions mean the same thing. The interface should use accepted common standards
2. Visibility	The user should not have to wonder where they are in the system, what they can do next or what has happened after and action
3. Match	The image of the system perceived by the users should match the model the users have about the system
4. Minimalist	This involves judging whether any extraneous information is a distraction and a slow-down
5. Memory	Users should not have to memorize a lot of information to carry out tasks. Memory load reduces users' capacity to carry out main task
6. Feedback	The system should provide feedback about the user's actions
7. Flexibility and efficiency	Users should be allowed to use shortcuts or tailor frequent actions for their own needs
8. Error message	The system should alert the users to potential errors. The messages should be clear and precise
9. Prevent errors	The system has mechanisms in place to prevent errors from occurring
10. Closure	The completion of a task is clearly indicated
11. Reversible actions	The system allows the user to easily backtrack
12. Language	The language should be presented in a form, easily understandable by the intended user
13. Control	The user should be able to leave an unwanted state easily
14. Documentation	The user should be provided with easily accessible help and documentation

[a] Heuristics and modified definitions were adapted from the work presented by Zhang et al. [11].

Once evaluators read the instruction manual and confirmed that they understood the instructions, they were given the evaluation booklet (see Appendix A) and asked to read the narrative on each page with the screen shot of WebCIS. When they arrived at a page of interest that is, those WebCIS screens that we were interested in evaluating, the experts were asked to review the page, determine if they observed any heuristic violations, and to rank the severity of the violations using ratings 1–4 (Table 1). Violations were defined as problems that might potentially interfere with the end users' ability to interact effectively with the system. We asked our experts to evaluate 18 out of 32 screen shots in the booklet. It was not necessary to keep the 0 severity rating since each expert was evaluating his or her own list of violations. As the experts conducted the modified heuristic evaluation, a member of the team guided them through the process and answered any questions they had.

Table 3
Reviewers choice of appropriate heuristics

Heuristic	1	2	3
1. Consistency	√	√	√
2. Visibility			
3. Match	√	√	√
4. Minimalist	√	√	√
5. Memory	√	√	√
6. Feedback			
7. Flexibility and efficiency			
8. Error message			
9. Prevent errors			√
10. Closure		√	
11. Reversible actions		√	
12. Language	√	√	√
13. Control		√	
14. Document		√	

3.2. Results

3.2.1. Relevant heuristics identified

As shown in Table 3, our three reviewers agreed that there were five heuristics, which would be appropriate to evaluate screen shots of a user interface, namely: consistency, match, minimalist, memory, and language. We instructed our evaluators to use these five heuristics to assess the screen shots.

3.2.2. Modified heuristic evaluation

On average, each evaluator took approximately 1 h to complete the evaluation of the booklet. All four of the evaluators commented on nine of the screen shots and three commented on an additional six screen shots. Altogether, our four evaluators made a total of 108 comments (16, 33, 31, and 28 comments each) on the design and layout of our screen shots. Eight of the comments were not classified because they were not assigned a heuristic. The majority of the problems (41%) identified in the design of our interfaces were due to violations of the consistency heuristic (Fig. 1).

Of the 100 violations assigned heuristics, there were 11 that were not rated according to their severity, 51% of the final 97 valid heuristics were rated as Type 2 problems, i.e., cosmetic and of low priority, and there was 22% each of Type 1 and Type 3 problems. Six violations (6%) were considered to be Type 4 violations (usability catastrophe). These were defined as problems that the expert felt should be remedied prior to the deployment of the final product. A Type 4 problem was identified when one of the evaluators reviewed a page containing an electrolyte report showing an electrolyte panel with the individual electrolytes presented using an abbreviated format. Selecting a link for one of the components (i.e., "Na," in our example) takes the user to a 'Questions of Interest' page, which lists a number of questions about "sodium." As our evaluator indicated,

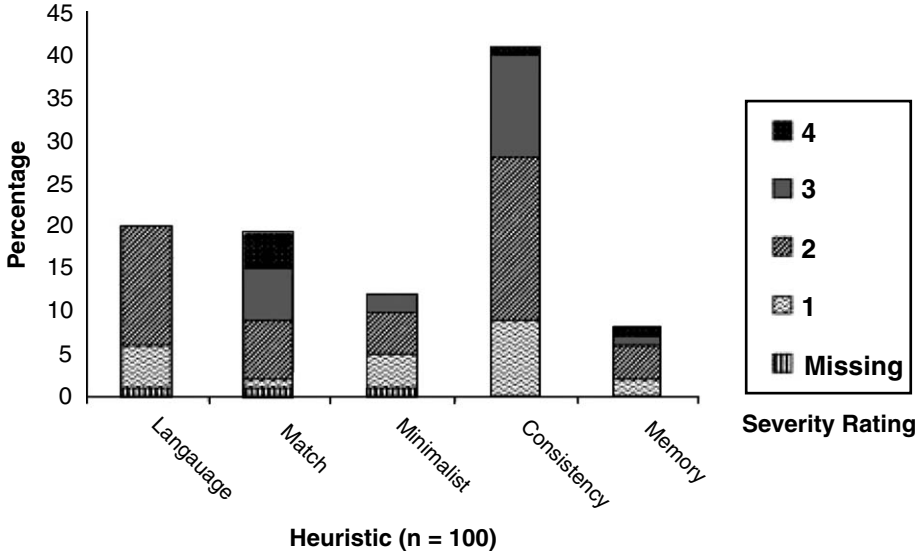

Fig. 1. Percentage of heuristic violations. This figure shows the percentage of each heuristic violation found ($n = 100$). For each heuristic, the severity rating frequency is shown. Each heuristic violation for which there is not a severity rating is categorized as missing.

Table 4
Technical solutions to usability problems

Technical	Definition	%[a]
Design	Refers to aspects of the layout of the page	43
Question	Refers to the questions on the 'Questions of Interest' page	33
Scope	Refers to the relationship of the concept of interest to the infobutton	16
WebCIS	Refers to a problem with WebCIS: this is outside the scope of the project	7
Specification	Refers to the need to change how the Infobutton Manager is called	1

[a] $n = 101$.

many clinicians will recognize the connection between "Na" and "sodium," but if this were a less common laboratory test or, worse, a medication that was unfamiliar to the user, the incongruency may lead to an error. The original Web page design provided the user with no clues to help him or her to recognize the connection between the initial word (Na) he selected and the subject of the page displayed subsequently, in this case sodium.

4. Technical solutions to heuristic violations

4.1. Method

In addition to conducting the evaluation, we wanted a descriptive method to characterize the solutions to the identified heuristic violations. These descriptions would assist in the implementation of the technical solutions to the violations once the Web page was built or modified. Three members of the team reviewed the results generated by the expert evaluators. The team members categorized the results based on the potential technical solutions that would have to be implemented to address the heuristic violations. Where two or more of the team members agreed, the violation was placed in that category.

4.2. Results

We analyzed the 108 comments and found that we could categorize the problems into five groups that could be addressed by the technical solutions as shown in Table 4. Of the 108 comments that were reviewed, there were seven comments that were not classified because team members could not agree on how they should be classified. Looking at the potential technical solutions, we found that most of our problems were due to the design or layout of the pages we had developed. In addition, we also found that there were issues with the questions themselves on the "Questions of Interest" page.

5. Resolution of violations

5.1. Examples

The following four screen shots[2] were included to depict examples of the heuristic violations and the solutions that were made to the pages before deployment.

5.1.1. Laboratory Results page: before HE

Fig. 2 shows a screen shot of a laboratory page in WebCIS. The evaluators were asked to review this page with respect to the information contained in the lower half of the screen. Selecting the follicle stimulating hormone (*FSH*) link in the upper frame results in the display of the FSH results in the lower frame. As shown in the figure, all rows are comments referring to the same test result. Our evaluators were concerned with the position of the infobutton (ⓘ) and the concept of interest—*FSH* (Fig. 2). This was coded as a problem with the "scope" of the infobutton. In addition, there were comments about the redundant infobuttons and *FSH* links. Further, the experts questioned whether selecting different *FSH* links or infobuttons in this context produces different results. Here, the notion of minimalist design was violated. In addition, the match heuristic was violated, as the evaluators' notion of what happened when each button was selected and what actually happened was different.

5.1.2. Laboratory Results page: after HE

Once we reviewed the comments made by our expert evaluators, we modified the Web page to reflect those comments. As shown in Fig. 3, we moved the infobutton icon so that it is closer to the concept of interest. In addition, we removed the multiple *FSH* hyperlinks and the associated redundant infobuttons. By implementing these solutions, we resolved the scope, minimalist, and match violations.

5.1.3. Questions of Interest page: before HE

Fig. 4 shows an example of the "Questions of Interest" page. The questions listed on this page provide links to resources that address the information needs events identified during our original observational study. When the user selects a link on this page, he or she is directed to a solution to that information need.

Selecting an infobutton, shown in Fig. 2, would open the window shown in Fig. 4. The evaluators felt that the naming of the concept of interest was inconsistent—the *FSH* link was selected and a Web page about the synonym *FOLLITROPIN* opens. This places an additional cognitive burden on the user, who has to remember that

[2] All screen shots show data from a fictional patient.

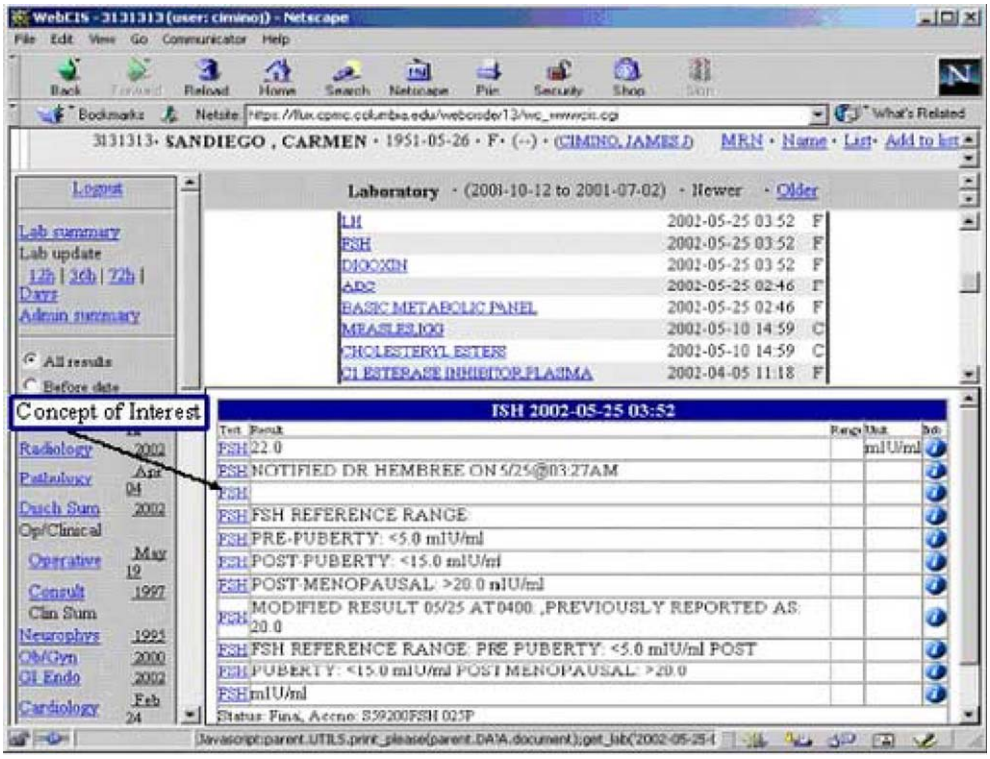

Fig. 2. Laboratory Results Web page—before HE. This screen shot shows a typical laboratory page in WebCIS before modifications were implemented. As shown, there are a number of hyperlinks and infobuttons (), which all refer to the same concept of interest—*FSH*. This interface violated the minimalist and match heuristics. In addition, there is discrepancy with the location of the infobutton and the concept of interest, i.e., the scope of the infobutton. Selecting the infobutton will take the user to the "Questions of Interest" page (Fig. 4).

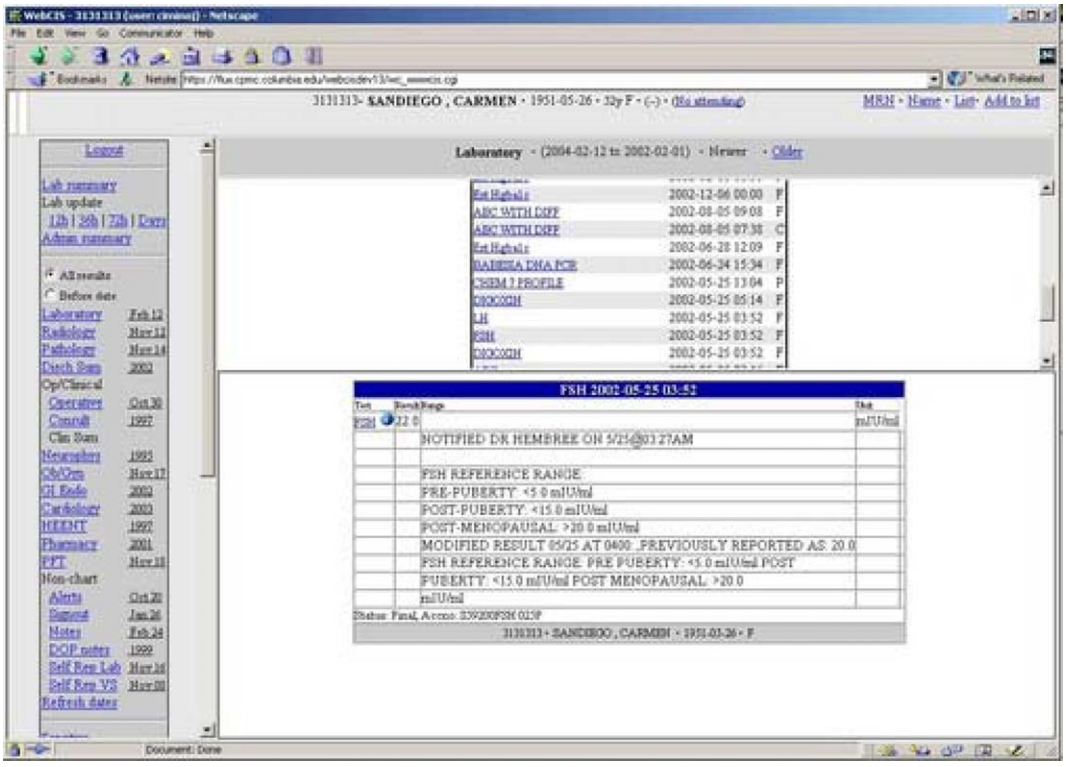

Fig. 3. Laboratory Results Web page—after HE. This screen shot shows the new laboratory page in WebCIS after the modifications were put into effect. The redundant links and infobuttons were removed. The concept of interest *FSH* and the infobutton are in closer proximity.

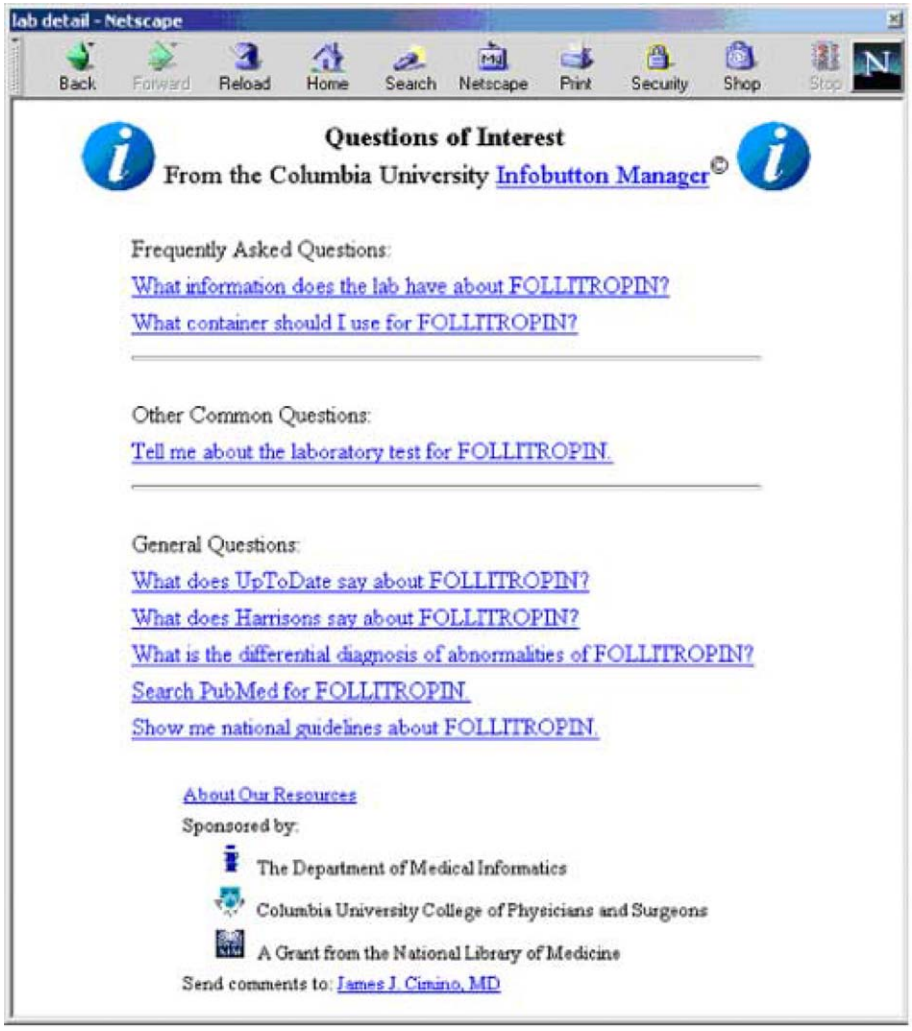

Fig. 4. Questions of Interest Web page—before HE. Selecting an infobutton would take the user to the "Question of Interest" page. The questions listed on this page provide links to resources that provide solutions to the user's information need. This page violated the consistency, memory, and minimalist heuristics. The technical solution to these violations involved addressing the design of the page and rewording the questions.

these terms are synonyms, thus violating the memory heuristic.

In addition, the minimalist heuristic was violated. In the original design of the "Questions of Interest" page the concept of interest, *FOLLITROPIN*, was placed at the end of each question, whereas a minimalist representation would seek to use this term only once. The solution to this problem was to place the concept of interest at the top of the page (Fig. 5).

Additional problems identified by the evaluators included question wording and general organization of information resources. For example, the evaluators indicated that the wording of a few of the questions was obtuse, e.g.: "What container should I use for FOLLITROPIN?" and that a few of the items on the "Questions of Interest" page where not questions, e.g., "Search PubMed for FOLLITROPIN." Furthermore, in the general questions section the user was forced to rely on his memory to know what the terms "Harrison's" or "UpToDate" referred to. There were no indicators to the user that allowed him to know what the resources provided prior to selection of the specific link. The technical solutions to these problems were to change the design of the page and reformat the questions displayed to the user.

5.1.4. Questions of Interest page: after HE

Fig. 5 shows a screen shot of the revised "Questions of Interest" page. Here, based on the usability experts' comments and the identified solutions, the concept of interest is linked directly to the term used to perform the search. In addition, the questions were streamlined to reflect the previous comments. The page is also sparser to reflect the minimalist design.

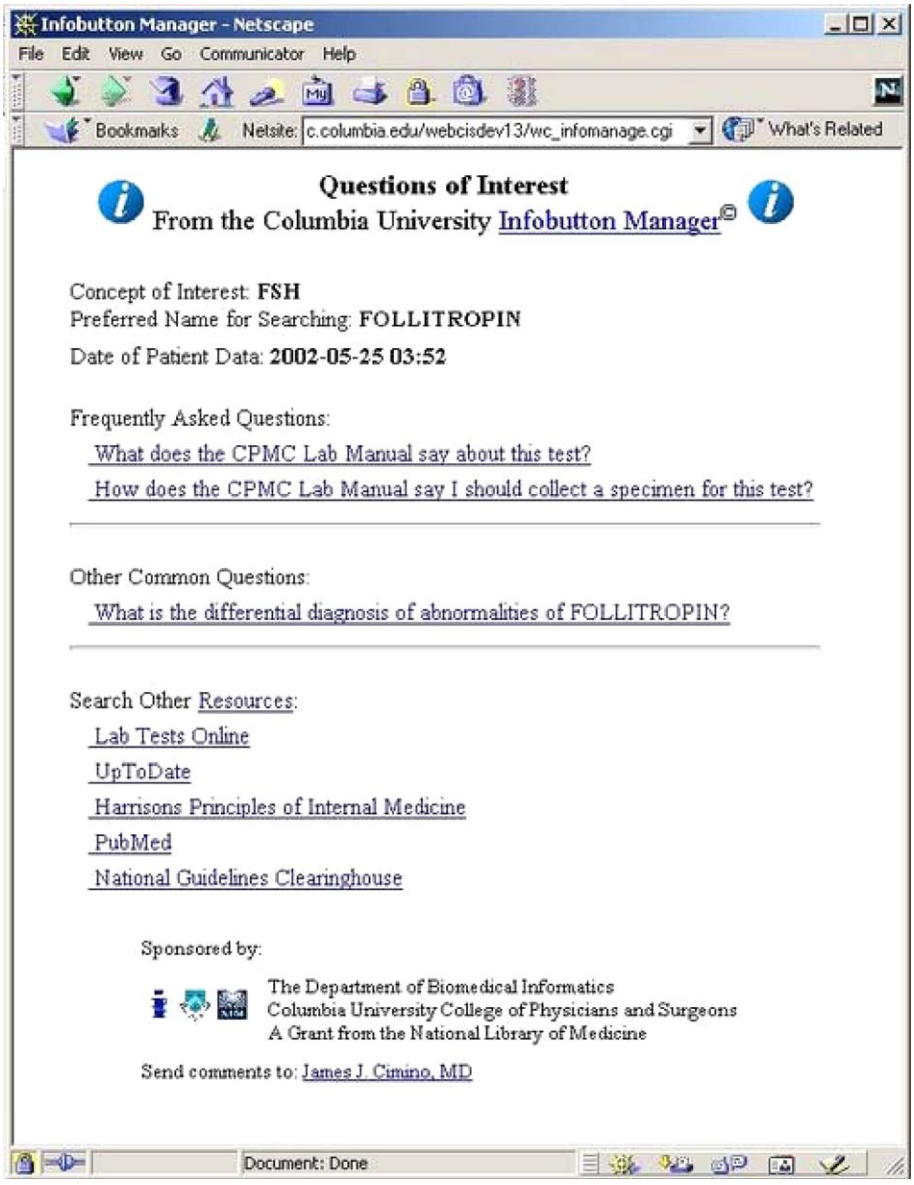

Fig. 5. Questions of Interest Web page—after HE. This screen shot shows the redesigned ''Questions of Interest'' page. The concept of interest and the term used to perform the match are linked directly. The questions were reworded to reflect the evaluates' comments.

5.2. Resolution of violations

A number of changes have been implemented in WebCIS drawing on the results of the paper-based heuristic evaluation. A resolution of a heuristic violation was said to have occurred if we were able to implement a technical solution to the violation. With this in mind, reviewing the solutions to heuristic violations, we found that we were able to address 70% of the violations. Fig. 6(top) shows the proportion of each heuristic that was addressed. It also shows the violations that were remedied based on the original perceived technical approach (middle), and severity of the violation (bottom). Of those instances where comments were related to technical aspects over which we had no control, we were unable to implement a remedy. Finally, appropriate solutions for a few Web pages, consistent with the previous solutions and the evaluates' comments, have eluded us.

6. Discussion

This paper describes an expeditious, straightforward, and inexpensive method to conduct an initial usability test of a user interface. Paper prototyping is not new [24,25], however, unlike many of the paper prototyping methods described, we chose to use screen shots of the user interface for the evaluation and not paper models of the user interface. To make the interaction more typical of the user's interaction with WebCIS we could have given the evaluators every permutation of Web pages

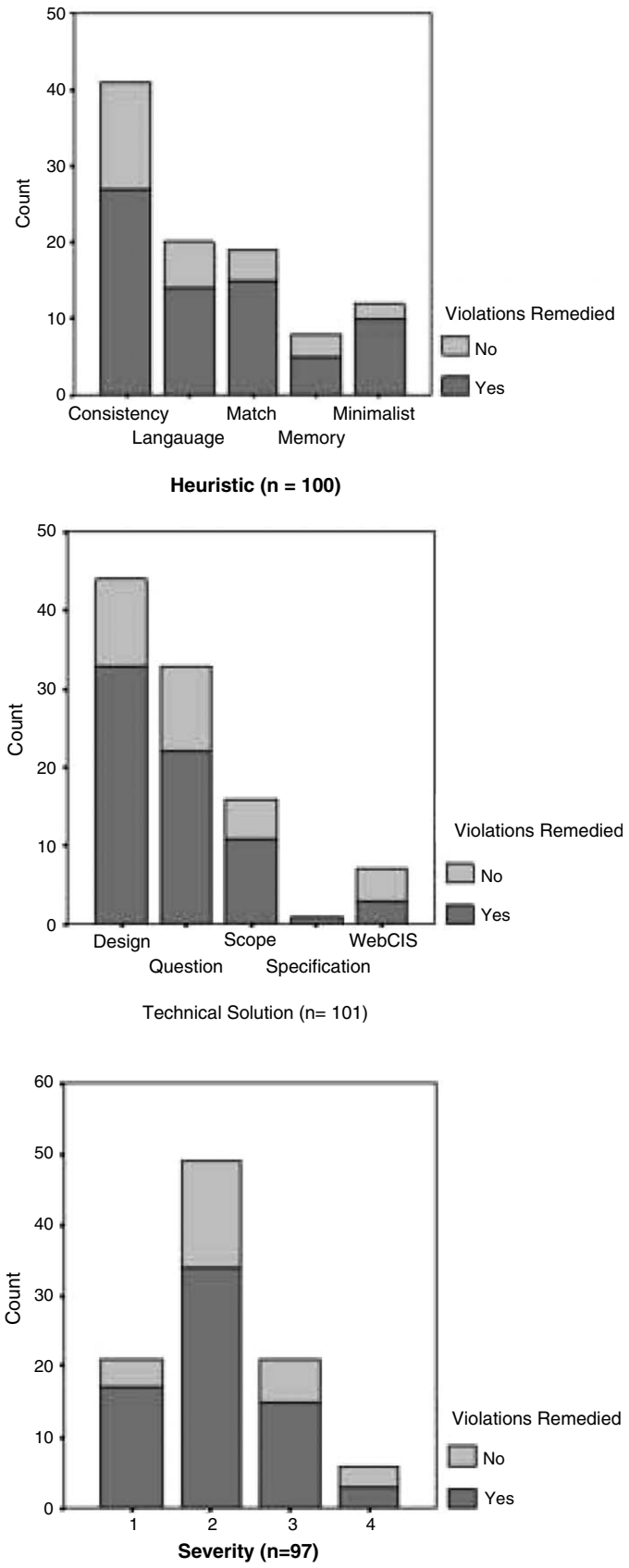

Fig. 6. Frequency of finding a remedy or solution based on the type of heuristic violation (top), the proposed technical solutions (middle), and the severity of the problem (bottom).

mimicking this interaction. However, we were limited in our ability to make changes so we opted to evaluate those pages of interest to us and not to evaluate the whole CIS.

Our study found that screen shots of a user interface could be used effectively by experts to conduct a heuristic evaluation. As we have shown, screen shots of the user interface can provide sufficient detail for a usability expert to utilize during this type evaluation and in fact, we implemented solutions for 70% of identified problems. It is possible to use current tools such as hypertext editors and graphic design programs, to create mock-ups of screen shots of the user interface even before the definitive interface is constructed. These mocked-up pages can be evaluated for potential usability problems early on, before substantial time and money are committed to the project.

However, this methodology is not without its limitations; the major one is its inability to assess functionality of the user interface. It does not take into account user interaction with the interface, or accessibility issues. Consequently, it is very difficult for the evaluator to have an understanding of how the user will ultimately interact with the system and whether the user interface violates any of the other nine heuristics (Table 2). For example, it is difficult for the evaluator to determine readily state transitions, adequacy of error messages, whether users can determine where they are in the system, and whether system response time is adequate. A second limitation of the study is that we did not ask the usability experts to assess if our solutions for the heuristic violations were adequate.

Similar to other studies [8], we found that we had both major and minor violations. As we reviewed the violations, it became obvious that we would not be able to address all of them due to the current design of WebCIS and the Infobutton Manager. A small proportion of the problems identified were related to the design of the CIS and although these issues were reported to the WebCIS team, the solution did not apply to the current project. Both the severity rating and technical mapping provided an efficient method by which we could determine which of the problems identified should, and could, be addressed.

The strength of this method lies in its ability to detect violation of Web page design and layout expeditiously, inexpensively, and in a straightforward manner. Screen shots of potential Web pages can be assessed *before* they are created, which may significantly reduce development costs related to the final interface with which the users will interact. Each expert took approximately an hour to complete the evaluation. Furthermore, asking the experts to rate the severity as they conducted the evaluation provided a method by which the need to convene all evaluators for a post hoc determination was eliminated and made the evaluation process even more expeditious. Instead, the research team collated the information, categorized the problems, and identified the solutions. Using this method makes it is possible to conduct end user testing more promptly using a functional prototype.

Combined with other methodologies this technique can be used to provide a richer assessment of the user interface design. Methods from the cognitive and usability engineering domains, i.e., cognitive task analysis, cognitive walkthrough, and formal user assessment might, in addition, complement the iterative design process of the user interfaces and might add to the assessment of the functionality of the interface.

Finally, there is an imperative to reduce medical errors [26] and to use technology to achieve this end. For that reason, as clinicians increasingly use clinical information systems, the user interface is of critical importance, since a poorly designed interface might itself become a potential source of medical errors. Formal assessment of interaction models, accessibility issues, and user experience may contribute to the detection of potential causes of medical error. However, preliminary assessment of interfaces for design and layout issues (before formal assessment) may also contribute in the reduction of potential errors. This study concentrated on a focused area of usability assessment, the paper-based evaluation, with the belief that even small changes, ultimately, may be beneficial to the end user. We feel that the strengths of this method outweigh its limitations especially if used in conjunction with other usability inspection methods. In addition, this paper sought to describe the simplification of an accepted usability testing methodology. Further, a comparison of the outcomes of this methodology with other usability techniques is necessary to facilitate validation of the methodology.

7. Conclusions

HE is a popular usability inspection method that uses experts to evaluate the design of user interfaces of devices. The paper-based HE methodology as outlined in this paper is complementary to the iterative, spiral information systems development process common in software engineering development life cycle. This modified HE method is simple to develop and implement, inexpensive, and expeditious. In addition, at a time when there is an obligation to reduce medical errors, it is imperative that the user interface with which the clinician interacts does not contribute to those errors. This method can be used to evaluate any type of user interface, in any domain, so long as it is possible to make paper-based screen shots of the interface.

Acknowledgments

The authors thank Justin Starren, William D. Roberts, Mark Graham, and Laura Slaughter for participating in this heuristic evaluation. We thank them for contributing their expertise to this project. This work is supported in part by NLM Grant 1R01LM07593 and NLM Training Grant LM07079-11.

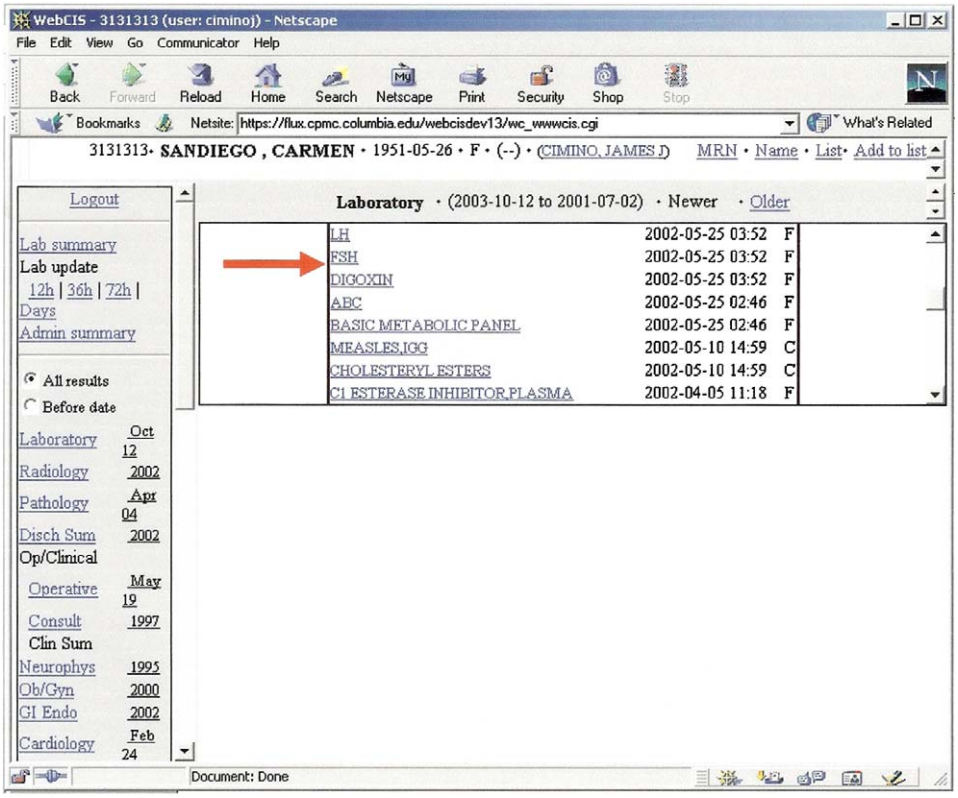

Fig. 7. A page taken from the evaluation booklet showing how an end user interacts with the system. A narrative is included to assist the experts as they conduct their evaluation.

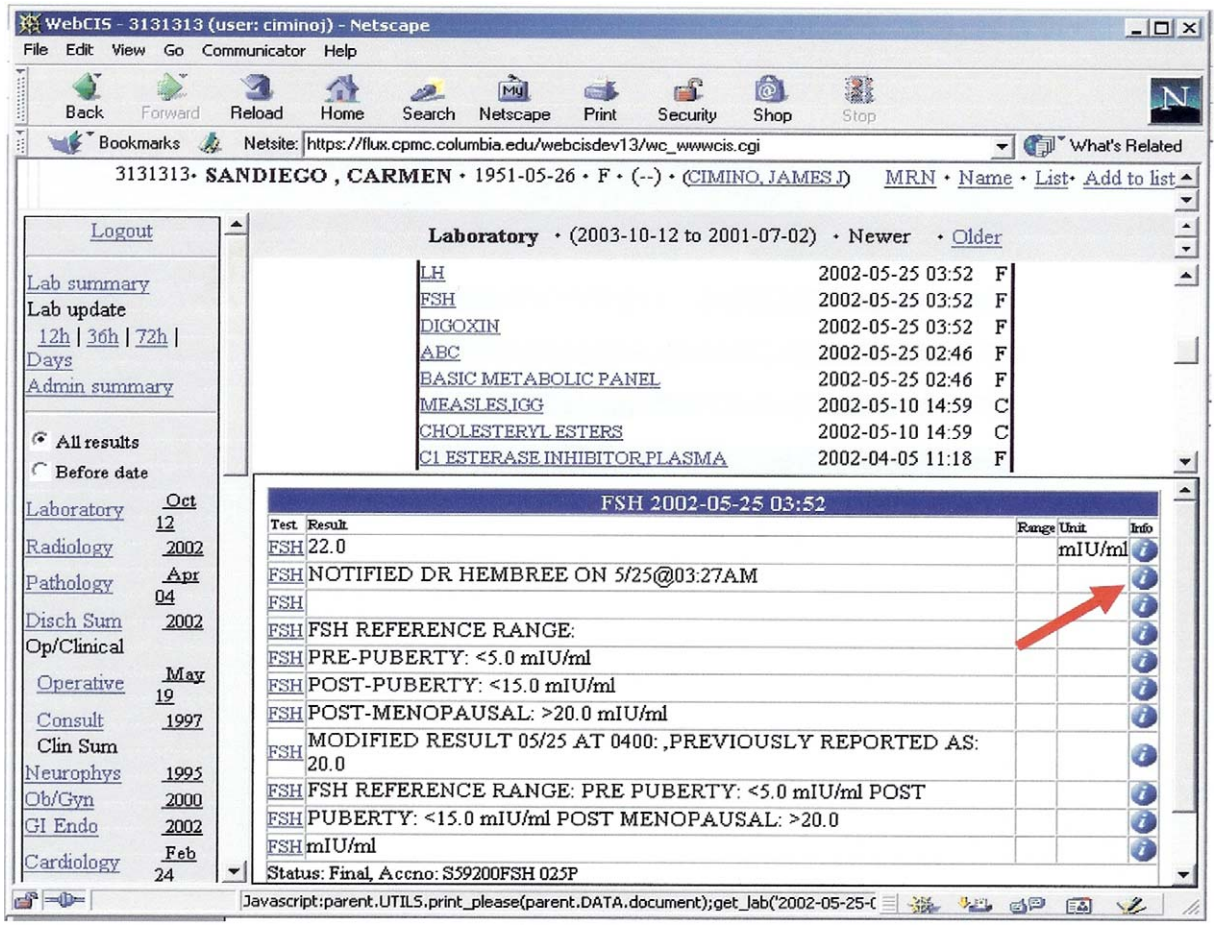

Fig. 8. This page follows Fig. 7. Here, the experts are asked to perform the HE of the page and to rate the severity of each violation.

Appendix A

Figs. 7 and 8 are screen shots of pages taken from the evaluation booklet that was given to the experts. These figures depict the user interface as an end user interacts with the system. The experts were asked to read the narration, and as shown in the second figure, once they arrived at a page of interested they were asked to conduct the HE at that stage.

References

[1] Szabo M, Kanuka H. Effects of violating screen design principles of balance, unity, and focus on recall learning, study time, and completion rates. J Educ Multimedia Hypermedia 1999;8(1):23–42.
[2] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005;293(10):1197–203.
[3] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004;11(2):104–12.
[4] Elkin P, Sorensen B, De Palo D, Poland G, Bailey K, Wood D, et al. Optimization of a research web environment for academic internal medicine faculty. J Am Med Inform Assoc 2002;5(5):472–8.
[5] Nielsen J. Finding usability problems through heuristic evaluation. In: Proceedings of the SIGCHI conference on human factors in computing systems. Monterey, California, USA: ACM Press; 1992. p. 373–80.
[6] Nielsen J, Molich R. Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI conference on human factors in computing systems: empowering people. Seattle, Washington, USA: ACM Press; 1990. p. 249–56.
[7] Beuscart-Zephir MC, Leroy N, Alao O, Darmoni S. Usability assessment study of a web site displaying medical resources online: the CISMeF. Stud Health Technol Inform 2002;90:133–7.
[8] Jeffries R, Miller JR, Wharton C, Uyeda K. User interface evaluation in the real world: a comparison of four techniques. New Orleans, Louisiana, USA: ACM Press; 1991.
[9] Nielsen J. Usability engineering. Boston: Academic Press; 1993.
[10] Graham MJ, Kubose TK, Jordan D, Zhang J, Johnson TR, Patel VL. Heuristic evaluation of infusion pumps: implications for patient safety in Intensive Care Units. Int J Med Inform 2004;73(11–12):771–9.
[11] Zhang J, Johnson TR, Patel VL, Paige DL, Kubose T. Using usability heuristics to evaluate patient safety of medical devices. J Biomed Inform 2003;36(1–2):23–30.
[12] Lathan CE, Sebrechts MM, Newman DJ, Doarn CR. Heuristic evaluation of a web-based interface for internet telemedicine. Telemed J 1999;5(2):177–85.
[13] Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 2004;37(1):56–76.
[14] Currie LM, Graham M, Allen M, Bakken S, Patel V, Cimino JJ. Clinical information needs in context: an observational study of clinicians while using a clinical information system. AMIA Annu Symp Proc 2003:190–4.
[15] Graham MJ, Currie LM, Allen M, Bakken S, Patel V, Cimino JJ. Characterizing information needs and cognitive processes during CIS use. AMIA Annu Symp Proc 2003:852.

[16] Cimino JJ, Li J, Graham M, Currie LM, Allen M, Bakken S, et al. Use of online resources while using a clinical information system. AMIA Annu Symp Proc 2003:175–9.
[17] Hughes J, Parkes S. Trends in the use of verbal protocol analysis in software engineering research. Behav Inform Technol 2003;22(2):127–40.
[18] Rieman J, Franzke M, Redmiles D. Usability evaluation with the cognitive walkthrough. In: Conference on human factors in computing systems; 1995. Denver, Colorado, United States; 1995. p. 387–8.
[19] Polson P, Lewis C, Rieman J, Wharton C. Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. Int J Man Mach Stud 1992;36(5):741–73.
[20] Nielsen J. Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI conference on human factors in computing systems: celebrating interdependence. Boston, Massachusetts, USA: ACM Press; 1994. p. 152–8.
[21] Shneiderman B. Designing the user interface. 3rd ed. Reading, MA: Addison-Wesley; 1998.
[22] Cimino JJ, Li J. Sharing infobuttons to resolve clinicians' information needs. AMIA Annu Symp Proc 2003:815.
[23] Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput 2003;35(3):379–83.
[24] Grady HM. Web site design: a case study in usability testing using paper prototypes. In: Proceedings of IEEE professional communication society international professional communication conference and Proceedings of the 18th annual ACM international conference on computer documentation: technology and teamwork; 2000. Cambridge, MA: IEEE Educational Activities Department; 2000. p. 39–45.
[25] Smith C. Transforming user-centered analysis into user interface: the design of new-generation products. In: Wood L, editor. User interface design bridging the gap from user requirements to design. Boca Raton, FL: CRC Press; 1998. p. 291–2.
[26] Kohn LT, Corrigan JM, Donaldson MS, editors. To err is human: building a safer health system. Washington, DC: The National Academies Press; 2000.