10-22-2007

Usability inspection methods after 15 years of research and practice

David G. Novick
University of Texas at El Paso, novick@utep.edu

Tasha Hollingsed
Lockheed Martin, tasha.hollingsed@lmco.com

Usability Inspection Methods after 15 Years of Research and Practice

Tasha Hollingsed
Lockheed Martin
2540 North Telshor Blvd., Suite C
Las Cruces, NM 88011
+1 505-525-5267
tasha.hollingsed@lmco.com

David G. Novick
Department of Computer Science
The University of Texas at El Paso
El Paso, TX 79968-0518
+1 915-747-5725
novick@utep.edu

ABSTRACT
Usability inspection methods, such as heuristic evaluation, the cognitive walkthrough, formal usability inspections, and the pluralistic usability walkthrough, were introduced fifteen years ago. Since then, these methods, analyses of their comparative effectiveness, and their use have evolved in different ways. In this paper, we track the fortunes of the methods and analyses, looking at which led to use and to further research, and which led to relative methodological dead ends. Heuristic evaluation and the cognitive walkthrough appear to be the most actively used and researched techniques. The pluralistic walkthrough remains a recognized technique, although not the subject of significant further study. Formal usability inspections appear to have been incorporated into other techniques or largely abandoned in practice. We conclude with lessons for practitioners and suggestions for future research.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Evaluation/methodology, training, help, and documentation.*

General Terms
Human Factors, Measurement

Keywords
Usability, evaluation, heuristic evaluation, cognitive walkthrough, pluralistic usability walkthrough, formal usability inspection

1. INTRODUCTION
By the early 1990s usability became a key issue, and methods for assuring usability burgeoned. A central argument in the field was the relative effectiveness of empirical usability testing versus other, less costly, methods (see, e.g., [10], [19], [20], [24]). Full-blown usability testing was effective but expensive. Other methods, generally known under the category of usability inspection methods [33], held the promise of usability results that kept costs low by relying on expert review or analysis of interfaces rather than by observing actual users empirically. Several different approaches to usability inspection were proposed, including heuristic evaluation [32], the cognitive walkthrough [51], the pluralistic walkthrough [3], and formal inspections [22].

At the peak of the trend toward development of usability inspection methods, Mack and Nielsen organized a workshop on the subject at the ACM CHI 92 conference [27]. In the fifteen years since the 1992 workshop, what happened to these methods? Did they further evolve? Is there evidence in the research literature of their use in practice?

After the workshop, Nielsen and Mack edited a book [33] that is the primary compilation of material about the movement toward usability inspection. In this paper, we track the fortunes of the methods and analyses presented in the book. We look at which of these methods and analyses led to use and to further research, and which led to relative methodological dead ends.

One of the chapters [54] was a historical review of usability methods, and citations to this work have been limited to subsequent historical reviews. Some chapters ([9], [23]) did not propose new usability inspection methods but rather compared the effectiveness of different approaches to usability evaluation; we discuss these contributions below. Other chapters offered perspective on situating usability approaches in the development process ([5], [6], [8], [18], [26], [52]). Citations, if any, to these contributions have tended to be limited to historical review of the development of usability evaluation rather than indicating evidence of further development of the approach or its use in practice.

To provide perspective on usability inspection methods in light of the 15 years of research and practice since the 1992 workshop, we review subsequently reported work for the four principal approaches described in [33]: heuristic evaluation, the cognitive walkthrough, formal usability inspections, and the pluralistic usability walkthrough. We conclude with some lessons for practitioners and suggestions for further research.

2. HEURISTIC EVALUATION
In 1990, Nielsen and Molich introduced a new method for evaluating user interfaces called heuristic evaluation [28], [31]. The method involves having a small group of usability experts evaluate a user interface using a set of guidelines and noting the severity of each usability problem and where it exists. They found that the aggregated results of five to ten evaluators of four interfaces

SIGDOC'07, October 22–24, 2007, El Paso, Texas, USA.

identified 55 to 90 percent of the known usability problems for those interfaces. They concluded that heuristic evaluation was a cheap and intuitive method for evaluating the user interface early in the design process. This method was proposed as a substitute for empirical user testing [32], [35].

After the publication of the initial papers on heuristic evaluation, other researchers began to compare the results of heuristic evaluation to the results of other methods. One study [19] compared the four best-known methods of usability assessment: empirical usability testing, heuristic evaluation, the cognitive walkthrough, and software guidelines. The study reported that heuristic evaluation found more problems than any other evaluation method, while usability testing revealed more severe problems, more recurring problems and more global problems than heuristic evaluation. Other researchers found that heuristic evaluation found more problems than a cognitive walkthrough, but only for expert evaluators [10]. System designers and non-experts found the same number of problems with both heuristic evaluation and cognitive walkthrough. Another study [24], however, found empirical testing to yield more severe problems than inspection methods and ascribed differences from the results found in [19] to differences in evaluator expertise.

To explain these differences in comparative results of usability assessment methods, Nielsen examined the role of expertise as a factor in the effectiveness of heuristic evaluation [32]. He compared evaluation results from three distinct groups of evaluators: novice, usability experts, and double experts, who have expertise both in usability and in the particular type of interface being evaluated. As could be expected, the double experts found more problems than the other two groups, but the two sets of experts reported the same number of problems in all areas other than the specific problems for the specific type of interfaces. Nielsen concluded that if only regular usability experts can be obtained, then a larger number of evaluators (between three and five) are needed than if double experts are available.

While these studies were intended to compare different evaluation methods and to highlight the advantages and disadvantages of each method, some took the results to conclude that heuristic evaluation was the best method to use in any case (e.g., [48]). Jeffries and Desurvire [20] addressed this misconception by clearly listing all of the disadvantages of heuristic evaluation found throughout these studies. A first disadvantage is that the evaluators must be experts, as suggested by Nielsen's findings earlier in the same year. A second disadvantage is that several evaluation experts are needed. The authors pointed out that it is difficult for some developers to obtain just one expert, much less many, and to use these experts several times throughout the development cycles, as Nielsen suggests, can become costly. A third disadvantage is the cost. Most of the issues identified by heuristic evaluation in the studies were minor, and few of the severe issues were identified. Another cost is that some of the issues may be "false alarms," issues that may never bother users in actual use.

Muller, Dayton, and Root [29] reviewed the previous three comparison studies and found that the methods used varied greatly and reanalyzed the data. They concluded that any decision made about the superiority of any method over another was premature. At the same conference, Nielsen responded to these comparison studies with one of his own [35]. He found that while user testing is more expensive than heuristic evaluation, it also provides a better prediction of the possible issues within the interface.

While the previous work stressed that the cost of conducting an inspection method is less than the cost of conducting an empirical user test, Wixon et al. [54] examined the overall cost of using inspection methods during development. They argued that the costs important to the overall commercial success of a product are not just within the costs of conducting the inspection or user testing, but also found in the costs of addressing the issues raised by the inspection. Given that inspections can be done earlier in the development process, the cost of addressing the issues found would be less than if the issues were found later in the process as with user testing.

The consensus view appears to be that empirical testing finds more severe issues that will likely impede the user, but at a greater cost to the developer. Heuristic evaluation finds many of the issues within the interface, but cheaper and earlier. At the same time, it is also possible to apply empirical usability testing to mock-up or other "low-tech" interfaces early in the design cycle [38].

Assessment of the effectiveness of heuristic evaluation continues as an active research thread. For example, an ethnographic study of paramedics using a prototype interface indicated that heuristic evaluation had disclosed some of the interface's usability problems, but that field observation of the interface in use revealed problems that were not identified in heuristic evaluation [47].

In recent years, heuristic evaluation has been used in a variety of projects, including development of a "seamless" system for video on demand [15] and several academic projects around the world (e.g., [30] and [43]).

3. COGNITIVE WALKTHROUGH

The cognitive walkthrough is a usability inspection method that evaluates the design of a user interface for its ease of exploratory learning, based on a cognitive model of learning and use [53]. Like other inspection methods, a cognitive walkthrough can be performed on an interface at any time during the development process, from the original mock-ups through the final release. The process of the cognitive walkthrough comprises a preparatory phase and an analysis phase. During the preparatory phase, the experimenters determine the interface to be used, its likely users, the task, and the actions to be taken during the task. During the analysis phase, the evaluators work through the four steps of human-computer interaction developed by Lewis and Polson [25]:

1. The user sets a goal to be completed within the system.
2. The user determines the currently available actions.
3. The user selects the action that they think will take them closer to their goal.
4. The user performs the action and evaluates the feedback given by the system.

Wharton et al. [53] surveyed the history of the cognitive walkthrough prior to 1994. At the time, the two main limitations of the cognitive walkthrough were the repetitiveness of filling out the forms and the limited range of problems the process found ([53], [51], [19]). A new method of cognitive walkthrough addressed these limitations by using small groups instead of individual evaluators and rotating the form filling within the group, evaluating the simple tasks first, and by keeping a record of all problems identified during the evaluation, not just those identified during the process ([53], [41]). However, this new process has only been described and informally studied with anecdotal evidence that the new process alleviates these two limitations. Other criticisms of the cognitive walkthrough were that it does not provide guidelines about what makes an action clearly available to a user and that it is not known what types of actions are considered by a broad range of users [51]). Franzke [13] showed that the labeling of the action is important to the clarity of the action and suggested that designers should consult the user population to determine the correct choice for labeling if they are unsure.

While most of the studies indicated that the cognitive walkthrough can be accomplished by evaluators with experience in the process, John and Packer [21] found that the method is learnable and usable for novices. However, novices may have difficulty in applying the method early in the design process and in choosing task scenarios. In another real-world application of the cognitive walkthrough, Spencer [46] applied it in a large software development company and found difficulties arising from the social constraints of the development team. For the process to work, the usability specialist leading the walkthrough needed to avoid design discussion and defensiveness of team members throughout the walkthrough. The process itself needed to be streamlined by collapsing the first three steps into one to cover broader problems and helping the developers not to perceive the process as a waste of time.

The process of using the four cognitive walkthrough steps could effectively demonstrate errors that users can make while using a safety-critical interface. When it was applied to a Bayesian Belief Net, the designer was enabled to determine what user actions were prone to result in error and what interface items were likely to influence these errors [14]. The cognitive walkthrough was likewise applied to operating procedures and their corresponding documentation. Evaluation of a cognitive walkthrough for operating procedures indicated that it could be used effectively and provided valuable feedback to the developers of the procedures and documentation [37]. The cognitive walkthrough, in junction with latent semantic analysis, could evaluate how well Web sites support users [4]. This approach identifies three different types of problem areas: domain vocabulary knowledge, meaningful sub-region and link labels, and conventions used to represent page elements. Evaluation of the approach found that it enabled evaluators to be effective in determining problems within a Web site for a broad range of users.

Since its inception, and with its refinements and extensions, the cognitive walkthrough has been shown to be an effective inspection method that can be applied not just by cognitive scientists and usability specialists but also by novice evaluators. However, the choice of task scenario can be difficult; if the scenario is not adequately described, the evaluation is not as effective. Even more "streamlined" versions of the cognitive walkthrough continue to be developed [46].

The cognitive walkthrough appears to be in continued use. Other recent applications of the cognitive walkthrough include the development of a multimodal tourist guide to Paris [1], an open-source, Web-based digital library for the management and dissemination of social science research data [17], and a PDA-based game for teaching science [43].

4. PLURALISTIC USABILITY WALKTHROUGH

The pluralistic usability walkthrough [3] adapted the traditional usability walkthrough to incorporate representative users, product developers, members of the product team, and usability experts in the process. It is defined by five characteristics:

1. Inclusion of representative users, product developers, and human factors professionals;
2. The application's screens are presented in the same order as they would appear to the user;
3. All participants are asked to assume the role of the user;
4. Participants write down what actions they, as users, would take for each screen before the group discusses the screens; and
5. When discussing each screen, the representative users speak first.

The developers of the pluralistic usability walkthrough identified both benefits and limitations of the approach. On the positive side, this approach offers feedback from users even if the interface is not fully developed, enables rapid iteration of the design cycle, and—because the users are directly involved—can result in "on-the-fly" redesign. On the negative side, the approach must be limited to representative rather than comprehensive user paths through the interface, and users who, at a particular step, did not choose the path the group will follow must "reset" their interaction with the interface.

The perspectives represented in the pluralistic usability walkthrough range along the axis of role in the development process. That is, the differences in expertise that users, developers and human-factors experts bring to the process stem from their roles as stakeholders in the process. It is possible to bring diversity of perspective along different axes than role. In particular, a walkthrough or heuristic evaluation could involve participants who instead differ with respect to their focus on aspects of user interfaces [55]. Evaluation of perspective-based usability inspection, with three experts each focusing on a different set of usability issues, found a 30 percent improvement for both individual inspectors and as a group. Moreover, the average number of problems found by each perspective inspector was higher than that of each heuristic inspector [55].

The pluralistic walkthrough appears to be in active use for assessing usability. It is included in the Usability Professionals Association draft body of knowledge [49]. Available reports indicate that pluralistic usability walkthrough is used in industry. For example, it was applied when upgrading a graphics program to Windows NT [11] and in evaluating the design of a multimedia learning environment based on the metaphor of interactive drama [42].

While some human factors experts continue to conduct usability walkthroughs that do not combine stakeholder perspectives (e.g., [7], [50], [44]), it seems likely that use of the pluralistic usability walkthrough is widespread but that teams do not refer to it as such in published reports. Rather, the pluralistic feature of the walkthrough may have become such a standard practice that it need not be mentioned.

5. FORMAL USABILITY INSPECTIONS

Formal usability inspection is a review by the interface designer and his or her peers of users' potential task performance [22]. Like the pluralistic usability walkthrough, this involves stepping through the user's task. However, because the reviewers consist of human factors experts, the review can be quicker, more thorough, and more technical than in the pluralistic walkthrough. The goal is to identify the maximum number of defects in the interface as efficiently as possible. The review process includes task performance models and heuristics, a variety of human-factors expertise, and defect detection within the framework of the software development lifecycle. Like the cognitive walkthrough, formal usability inspections require definitions of user profiles and task scenarios. And, like the cognitive walkthrough, the reviewers use a cognitive model of task performance, which can be extended with a checklist of cognitive steps similar to those invoked by Norman [36] to bridge the “gulf of execution.”

Hewlett Packard used this method for at least two years before 1995. The inspection team included design engineers, usability engineers, customer support engineers, and customers at times. The team inspected fourteen products and found an average of 76 usability concerns per product and an average of 74 percent of those concerns were fixed per product. While no formal evaluation of the results was done, it was found that the engineers could detect several of the usability concerns, and the engineers enjoyed using the method while increasing their awareness of user needs [15].

Digital Equipment Corporation also conducted a version of formal usability inspections from 1994 to 1995 on ten products. They found an average of 66 usability problems per product and fixed an average of 52 problems per product. Finding even small usability problems proved to be an asset, especially when a number of these problems were easily fixed. As more problems were fixed, the perceived quality of the product improved as well, even if most of these fixes were small [45].

Since then, it appears that little research has been conducted on formal usability inspections. This approach now tends to be grouped into the overall inspection method class and gets overshadowed by the better known heuristic evaluation when comparisons between inspection and empirical methods have been conducted. As a method, formal usability inspection gains speed at the cost of losing the multiple stakeholder perspectives of the pluralistic walkthrough, and its cognitive model can be seen as less comprehensive than that of the cognitive walkthrough. Nevertheless, at one time the formal usability inspection was used in practice in at least two companies. In one of these cases, the design engineers on the inspection teams were able to find usability problems, to describe and implement fixes for them, and to gain greater awareness of the usefulness of usability inspections [22]. But the absence of literature indicating current use or comparing formal usability inspections to other approaches makes it difficult to conclude that formal usability inspections currently are as effective as other methods.

6. CONCLUSION

Both empirical usability testing and usability inspection methods appear to be in wide use, with developers choosing the most appropriate method for their purposes and their context. For example, half of the ten intranets winning a 2005 competition used heuristic evaluation [34]. The same report indicated that empirical usability testing was used by 80 percent of the winning intranets. The cognitive walkthrough appears to be in continued use, although reports of use are not as frequent. The pluralistic usability walkthrough remains in the repertoire of usability experts, although usability experts continue to conduct user-only usability walkthroughs. And formal usability inspection, although shown to be an effective approach for identifying usability problems, appears to be used less now than in the mid-1990s.

Many have claimed that usability inspection methods make for faster and more cost-efficient evaluation of the usability of an interface than empirical user testing. But while usability inspection methods do identify a number of usability problems faster and more cost-efficiently, the best performing evaluator and method still only found 44 percent of the usability problems found in a laboratory setting [9]. While the cognitive walkthrough is useful for predicting problems on a given task and heuristic evaluation is useful for predicting problems on the interface, empirical testing provides lots of information throughout the interface and is the benchmark against which all other methods are measured [9]. Indeed, Jeffries et al. [19] noted that evaluators of usability methods may have rated problems found though empirical usability testing as, on average, more severe precisely because the problems were identified empirically rather than analytically. While inspection methods need expert evaluators to be effective, their strengths are that they can be implemented into the early stages of the development cycle and provide a forum in which changes to interface can be discussed.

Fifteen years after the emergence of usability inspection methods, comparisons between inspection methods and empirical testing continue (e.g., [12], [55]), and comparisons of techniques for empirical usability testing continue, as well [2]. For example, in-depth analysis of the empirical usability testing suggested that the types of problems detected in these tests can be categorized by severity (prevents task, significant delay, minor effect, and suggestion) and scope (global and local) and then traced back to its source (tutorial, exercise accuracy, exercise efficiency, questionnaire, and verbal protocol) [12]. The verbal protocol (akin to the think-aloud method) found most of the problems within the testing (61 out of 65), but a large proportion of them (40 out of 65) were found only in the verbal protocol. The verbal protocol found problems in all severity levels, but most were less critical. The other four sources found most of the critical problems within the interface. These results suggested that as a benchmark when using inspection methods, usability engineers compare a verbal protocol results to the chosen inspection method, given that the verbal protocol can find numerous problems through all scopes and severities and requires less time and lower cost.

The research on comparisons of usability assessment methods suggests several lessons for practitioners. First, while “faster, cheaper” methods such as heuristic evaluation and the pluralistic usability walkthrough can be useful for rapid iteration early in the design cycle, inspection methods cannot fully substitute for the empirical user testing needed before releasing an interface or Web site to the public. Second, empirical methods can also be used early in the development process, via “low-tech” versions of interfaces. Third, developers often combine multiple inspection methods—heuristic evaluation and the cognitive walkthrough—in the same project so that they obtain better coverage of usability issues. And fourth, adding multiple perspectives—along dimensions such as the range of stakeholders or kinds of usability problems—appears to improve the effectiveness of inspection methods.

With usability inspection methods solidly represented in the methodological repertoires of usability professionals (see, e.g., [34], [39], [49], [50]), research issues are shifting from showing efficacy of the methods toward further adaptation of the methods as interaction media evolve (e.g., [4]). As new modalities such as video on cellular telephones become widespread, such adaptations become correspondingly urgent.

It remains an open issue as to why usability professionals, in practice, rely on single-perspective methods, typically involving users, or experts, but not both. The evidence from reports of recent uses of heuristic evaluation suggests that many usability specialists are missing the benefits of the pluralistic walkthrough and perspective-based evaluation. A further research question along these lines involves possible means of combining multiple dimensions of perspectives in a unified usability inspection method. That is, could one cross the pluralistic perspectives of different stakeholders with the differentiated technical foci of perspective-based heuristic evaluation? This may not be straightforward, as perspective-based heuristic evaluation relies on human-factors expertise, focused onto particular kinds of problems with interfaces, while the pluralistic approach, by definition, involves non-experts. Perhaps it would be possible, though, to integrate the non-expert stakeholders into the focused teams.

These research issues revolve around finding increasingly effective means of predicting and finding usability problems in interface designs spanning the development process. This defect-finding approach can be complemented by research into ways of lowering users' frustration. For example, one can look how to help users develop new ways of learning applications, and avoid frustration episodes because they have a greater knowledge of the application and its interface [40]. At a deeper level, though, a new direction for research should complement these defect-tracking and learning approaches by seeking to understand the root causes of usability problems. The ideal solution would be to know the reasons for usability problems, so that designers can minimize the effort spent on usability inspection and testing.

7. ACKNOWLEDGMENTS

The authors thank Karen Ward and Nigel Ward for their valuable comments.

8. REFERENCES

[1] Almeida, L., Amdal, I., Beires, N. , Boualem, M., Boves, L., den Os, E., Filoche, P., Gomes, R., Knudsen, J., Kvale, K., Rugelbak, J., Tallec, C., Warakagoda, N. (2002). The MUST guide to Paris: Implementation and expert evaluation of a multimodal tourist guide to Paris, *ISCA Tutorial and Research Workshop (IDS'2002)*, Kloster Irsee, Germany, June, 2002, available at http:// citeseer. ist. psu. edu/ cache/ papers /cs/ 26307/ http: zSzzSzlands. let. kun. nlzSzliteraturezSzboves. 2002. 1. pdf/ almeida02must. pdf, accessed May 26, 2007.

[2] Andreasen, M. S., Nielsen, H., Schrøder, S., and Stage, J. 2007. What happened to remote usability testing?: An empirical study of three methods, *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI 2007)*, San Jose, CA, April 28-May 3, 2007, 1405-1414.

[3] Bias, R. G. 1994. The pluralistic usability walkthrough: coordinated empathies. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 63-76.

[4] Blackmon, M., Polson, P., Kitajima, M., and Lewis, C. (2002). Cognitive walkthrough for the Web, *Proceedings of the Conference on Human Factors in Computing Systems (CHI 2002)*, Minneapolis, MN, April 20-25, 2002, 463-470.

[5] Blatt, L. A. and Knutson, J. F. 1994. Interface design guidance systems. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 351-384.

[6] Bradford, J. (1994). Evaluating high-level design: synergistic use of inspection and usability methods for evaluating early software designs. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 235-253.

[7] Brodie, C., Karat, C., and Karat, J. (2006). An empirical study of natural language parsing of privacy policy rules using the SPARCLE policy workbench, *Proceedings of the Second Symposium on Usable Privacy and Security (SOUPS '06)*, Pittsburgh, PA, July 12-14, 2006, 8-19.

[8] Brooks, P. (1994). Adding value to usability testing. . In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 255-271.

[9] Desurvire, H. (1994). Faster, cheaper!! Are usability inspection methods as effective as empirical testing? In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 173-202.

[10] Desurvire, H., Kondziela, J., and Atwood, M. (1992). What is gained and lost when using evaluation methods other than empirical testing, *Proceedings of the Conference on Human Factors in Computing Systems (CHI 92)*, Monterey, California, May 3-7, 1992, 125-126.

[11] Dunbar, A. (1997). CADNET design team asks customers: How does it work on paper? *The Culpepper eBulletin Newsletter* 157, available at http://www.culpepper.com/ eBulletin/1997/157cs.asp, accessed May 26, 2007.

[12] Ebling, M. and John, B. (2000). On the contributions of different empirical data in usability testing, in *Proceedings of the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, New York, NY, August 17-19, 2000, 289-286.

[13] Franzke, M. (1995). Turning research into practice: characteristics of display-based interaction, in *Proceedings of the Conference on Human Factors in Computing Systems (CHI 95)*, Denver, Colorado, May 7-11, 1995, 421-428.

[14] Galliers, J., Sutcliffe, A., and Minocha, S. (1999). An impact analysis method for safety-critical user interface design, *ACM Transactions on Computer-Human Interaction* 6(4), 341-369.

[15] Grasland, I., Kerbiriou, P., Janse, M., Issarny, V., Sacchetti, D., Talamona, A., Leal, A. (2004). *User experience report of the three demonstrators and the external application of the Ozone framework*, Ozone Project, available at http://www.hitech-projects.com/euprojects/ozone/public_ docs/del_oct04/ozone-tho-31aug04-d15d-pk.pdf, accessed May 27, 2007.

[16] Gunn, C. (1995). An example of formal usability inspections in practice at Hewlett-Packard company, *Proceeding of the Conference on Human Factors in Computing System (CHI 95)*, Denver, Colorado, May 7-11, 1995, 103-104.

[17] Hovater, J., Krot, M., Kiskis, D. L., Holland, H., & Altman, M. (2002). Usability testing of the Virtual Data Center, Workshop on Usability of Digital Libraries, *Second ACM-*

IEEE Joint Conference on Digital Libraries, Portland, OR, July 14-18, 2002, available at http: //www.uclic. ucl.ac.uk/annb/DLUsability/Hovater7.pdf, accessed May 26, 2007.

[18] Jeffries, R. (1994). Usability problem reports: helping evaluators communicate effectively with developers. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 273-294.

[19] Jeffries, R., Miller, J., Wharton, C., and Uyeda, K. (1991). User interface evaluation in the real world: a comparison of four techniques, *Proceeding of the Conference on Human Factors in Computing System (CHI 91)*, New Orleans, LA, April 27-May 2, 1991, 119-124.

[20] Jeffries, R., and Desurvire, H. (1992). Usability testing vs. heuristic evaluation: was there a contest? *ACM SIGCHI Bulletin* 24(4), 39-41.

[21] John B., and Packer, H (1995). Learning and using the cognitive walkthrough method: A case study approach, *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI 95)*, Denver, Colorado, May 7-11, 1995, 429-436.

[22] Kahn, M., and Prail, A. (1994). Formal usability inspections, in Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 141-171.

[23] Karat, C.-M 1994. A comparison of user interface evaluation methods, in Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 203-233.

[24] Karat, C.-M., Campbell, R. and Fiegel, T. (1992). Comparison of empirical testing and walkthrough methods in user interface evaluation, *Proceedings of the Conference on Human Factors in Computing Systems (CHI 92)*, Monterey, CA, May 3-7, 1992, 397-404.

[25] Lewis, C., and Polson, P. (1991). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces (tutorial), *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI 91)*, April 27-May 2, 1991, New Orleans, LA.

[26] Mack, R. and Montaniz, F. (1994). Observing, predicting, and analyzing usability problems. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 295-339.

[27] Mack, R., and Nielsen, J. (1993). Usability inspection methods: Report on a workshop held at CHI'92, Monterey, CA, May 3–4, 1992, *ACM SIGCHI Bulletin* 25(1) 28-33.

[28] Molich, R., and Nielsen, J. (1990). Improving a human-computer dialogue, *Communications of the ACM* 33(3), 338-348.

[29] Muller, M., Dayton, T., and Root, R. (1993). Comparing studies that compare usability assessment methods: an unsuccessful search for stable criteria, *INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems*, Amsterdam, April 24-29, 1993, 185-186.

[30] Naseem, A. (2005). Dynamics of user needs analysis in redesigning an open learning website: A case from Pakistan, *Proceedings of ASCLITE 2005*, Brisbane, Australia, December 4-7, 2005, 503-507.

[31] Nielsen, J., and Molich, R. (1990). Heuristic evaluation of user interfaces, *Proceedings of the Conference on Human Factors in Computing Systems (CHI 90)*, Seattle, WA, April 1-5, 1990, 249-256.

[32] Nielsen, J. (1992). Finding usability problems through heuristic evaluation, *Proceedings of the Conference on Human Factors in Computing System (CHI 92)*, Monterey, CA, May 3-7, 1992, 373-380.

[33] Nielsen, J. and Mack, R. (eds.) (1994), *Usability inspection methods*, John Wiley & Sons, Inc., New York.

[34] Nielsen, J. (2005). Ten best intranets of 2005, Jakob Nielsen's Alertbox, February 28, 2005, available at http: //www.useit.com/alertbox/20050228.html, accessed May 26, 2007.

[35] Nielsen, J., and Phillips, V, (1993). Estimating the relative usability of two interfaces: heuristic, formal, and empirical methods compared, *Proceedings of the Conference on Human Factors in Computing System (CHI 93)*, Amsterdam, April 24-29, 1993, 214-221.

[36] Norman, Donald A. (1988). *The design of everyday things*. Doubleday, New York.

[37] Novick, D. and Chater, M. (1999). Evaluating the design of human-machine cooperation: The cognitive walkthrough for operating procedures, *Proceedings of the Conference on Cognitive Science Approaches to Process Control (CSAPC 99)*, Villeneuve d'Ascq, FR, September 21-24, 1999.

[38] Novick, D. (2000). Testing documentation with “low-tech” simulation, *Proceedings of IPCC/SIGDOC 2000*, Cambridge, MA, September 24-27, 2000, 55-68.

[39] Perceptive Sciences (2006). Project briefs, available at http: //www.perceptivesciences.com/clients/project_briefs.php, accessed May 27, 2007.

[40] Rieman, J. (1996). A field study of exploratory learning strategies, *ACM Transactions on Computer-Human Interaction* 3(3), 189-218.

[41] Rieman, J., Franzke, M., and Redmiles, D. (1995). Usability evaluation with the cognitive walkthrough, *Proceedings of the Conference on Human Factors in Computing Systems (CHI 95)*, Denver, Colorado, May 7-11, 1995, 387-388.

[42] Rogers, T., Jagodzinski, A., Phillips, M., and Turley, S. (1997). Interactive drama: A tool for humanistic learning. *Proceedings of Computer Aided Learning (CAL) 97*, Exeter, UK, March 23-26, 1997, 249-257.

[43] Sánchez, J., Salinas, A., and Sáenz, M. (2006). Mobile game-based science learning, *Proceedings of the Conference on Distance Learning and the Internet (APRU DLI 2006)*, Tokyo, November 10-8, 2006, 18-30.

[44] Spagnolli, A., Gamberini, L., Cottone, P., and Mantovani, G. (2004). Ergonomics of virtual environments for clinical use. In Riva, G., Botella, C., Légeron, P., and Optale G., (eds.), *Cybertherapy: Internet and Virtual Reality as*

Assessment and Rehabilitation Tools for Clinical Psychology and Neuroscience, IOS Press, Amsterdam, 217-230.

[45] Sawyer, P., Flanders, A., and Wixon, D. (1996). Making a difference—the impact of inspections, *Proceedings of the Conference on Human Factors in Computing System (CHI 96)*, Vancouver, BC April 13-18, 1996, 376-382.

[46] Spencer, R. (2000). The streamlined cognitive walkthrough method, *Proceedings of the Conference on Human Factors in Computing System (CHI 2000)*, The Hague, The Netherlands, April 1-6, 2000, 353-359.

[47] Tang, Z., Johnson, T., Tindall, R., and Zhang, J. (2006) Applying heuristic evaluation to improve the usability of a telemedicine system, *Telemedecine Journal and E-Health* 12(1), 24-34.

[48] Tognazzini, B. (1992). *Tog on interface*. Addison-Wesley Publishing Company, Reading, MA.

[49] Usability Professionals Association (undated). Methods: Pluralistic usability walkthrough, available at http://www.usabilitybok.org/methods/p2049, accessed May 26, 2007.

[50] User-Centered Web Effective Business Solutions (2007). Heuristic usability evaluation, available at http://www.u-cwebs.com/usability/web-site-usability.htm, accessed May 27, 2007.

[51] Wharton, C., Bradford, J., Jeffries, R., and Franzke, M. (1992). Applying cognitive walkthroughs to more complex user interfaces: Experiences, issues, and recommendations, *Proceedings of the Conference on Human Factors in Computing System (CHI 92)*, Monterey, CA, May 3-7, 1992, 381-388.

[52] Wharton, C. and Lewis, C. 1994. The role of psychological theory in usability inspection methods. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 341-350

[53] Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994). The Cognitive Walkthrough Method: A Practitioner's Guide. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 1994, 105-140.

[54] Wixon, D., Jones, S., Tse, L., and Casaday, G. (1994). Inspections and Design Reviews: Framework, History and Reflection. In Nielsen, J. and Mack, R. (eds.), *Usability inspection methods*, John Wiley & Sons, Inc., New York, 1994, 77-103.

[55] Zhang, Z., Basili, V., Shneiderman, B. (1998). An empirical study of perspective-based usability inspection, *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, Santa Monica, CA, October 5-9, 1998, 1346-1350.