International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 10 | Oct - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

A Methodical Review of Usability Evaluation in Web Development

Khushboo Rinayat, Sanjana Choure, Subodhini Gupta

Sam Global University, Bhopal, Madhya Pradesh, India
khushburinayat0009@gmail.com
sanjanachoure06@gmail.com
subodhinig@gmail.com

ABSTRACT

The challenge of developing further usable Web operations has motivated the appearance of several ways, styles, and tools to address Web usability issues. Although there are numerous proffers for supporting the development of usable Web operations, numerous inventors aren't apprehensive of them and innumerable associations don't duly apply them. This paper reports on a systematic review of usability evaluation styles in Web development. The idea of the review is to probe what usability evaluation styles have been employed by experimenters to estimate Web vestiges and how they were employed. An aggregate of 51 exploration papers has been reviewed from an original set of 410 papers. The results show that 45 of the papers reviewed reported the use of evaluation styles specifically drafted for the Web and that the most employed system is stoner testing. In addition, the results of the review have linked several exploration gaps. Specifically, 80 of the evaluations are still performed at the perpetration phase of Web operations development and 47 of the papers didn't present any confirmation of the usability evaluation system(s) employed.

Keywords: Usability Evaluation Methods, Web development, Systematic Review.

I. INTRODUCTION

Usability is a pivotal factor in Web operation development. The ease or difficulty that druggies witness with systems of this kind will determine their success or failure. As Web operations have become the backbone of business and information exchange, the need for usability evaluation styles specifically drafted for the Web – and technologies that support the usability design process – has become critical. The challenge of developing further usable Web operations has motivated the appearance of various ways, styles, and tools to address Web usability issues. Although there are numerous proffers for supporting the development of usable Web operations, numerous inventors aren't apprehensive of them and innumerable associations do not duly apply them. To address this issue, several studies aimed at comparing usability evaluation styles for Web development were reported. These studies frequently compare a reduced number of evaluation styles, and the selection of styles is typically driven by the prospects of the experimenter. Thus, there's a need to identify, more methodically, what usability evaluation styles have been successfully applied to Web development. In this paper, we present a methodical review for assessing what usability evaluation styles have been employed for Web usability evaluation and their relation to the Web development process. Methodical reviews are useful for recapitulating all

information about a miracle of interest (e.g., a particular exploration question) in an unprejudiced manner. The thing of our review is, thus, to examine the current use of usability evaluation styles in Web development from the point of view of the following exploration questions what usability evaluation styles have been employed by experimenters to estimate Web vestiges and how were they employed? This paper is organized as follows. Section 2 discusses affiliated work. Section 3 presents the protocol we used to review the usability evaluation styles employed in Web development. Section 4 describes the results of the methodical review. Section 5 discusses the pitfalls to the validity of the results. Eventually, section 6 presents our conclusions and suggests areas for further disquisition.

II. Related Work

Several inquiries leveled at likening usability evaluation styles for trap evolution have been reported in the last many times one of the most comprehensive inquiries was published by Ivory and Hearst in 2002. They proffered a taxonomy for categorizing automated usability evaluation styles. The taxonomy was applied to 128 usability evaluation styles, where 58 of them can trap stoner interfaces. The effects of this check suggest encouraging ways to expand styles to support automated usability evaluation. Another study by Alva et al. offered an evaluation of seven styles and tools for usability evaluation in software productions and vestiges for the trap. The purpose of this study was to determine the place of community among the styles utilizing the principles outlined in the ISO 9241- 11metric. Still, this is an informal check with no outlined exploration questions and no hunt process to identify the styles that were considered. Batra and Bishu reported the effects attained with two usability evaluation inquiries for trap operations. The idea of the first study was to analogize the effectiveness and forcefulness between stoner testing and heuristic evaluation. The results showed that both styles managed veritably nonidentical usability cases and are inversely effective and operative for trap usability evaluation. The idea of the alternate study was to analogize the interpretation between remote and traditional usability testing. The effects indicate that there's no significant disparity between the two styles. Although several comparisons of usability evaluation styles have been reported, we aren't apprehensive about any methodical review published in the field of trap usability. The maturity of the published inquiries is informal literature checks or comparisons with no outlined exploration questions, no hunt process, no outlined data birth, or no data dissection process. We only set up two methodical reviews conducted in affiliated fields, Freire offered a methodical review on trap availability to identify ways for developing popular content in trap operations. This review includes 53 inquiries, and it also proposes a bracket of these ways tallying to the processes described in the ISO/ IEC 12207 standard issue. Mendes offered a methodical review to probe the rigor of calls of trap engineering exploration.

III. RESEARCH METHOD

A methodical review is a means of assessing and interpreting all accessible exploration that applies to a personal exploration question, content area, or miracle of interest. It aims to carry out a fair evaluation of exploration content by utilizing a secure, rigid, and auditable methodology. A methodical review involves several stages and conditioning. In calculating the review stage, the want for the review is linked, the exploration questions are prescribed, and the review protocol is outlined. In the conducting the review stage, the primary inquiries are named, the quality valuation exercised to carry inquiries is outlined, the data birth and monitoring are performed, and the attained data is synthesized. Eventually, in the reciting the review stage, the dispersion mechanisms are prescribed, and the review report is offered. The conditioning concerning the planning and conducting of our methodical review are described in the following subsections. The reporting of the review stage is offered in Section 4. A methodical review is a means of assessing and interpreting all accessible exploration that applies to a personal exploration question, content area, or miracle of interest. It aims to carry out a fair evaluation of exploration content by utilizing a secure, rigid, and auditable methodology. A methodical review involves several stages and conditioning. In calculating the review stage, the want for the review is linked, the exploration questions are prescribed, and the review protocol is outlined. In the conducting the review stage, the primary inquiries are named, the quality valuation exercised to carry inquiries is

outlined, the data birth and monitoring are performed, and the attained data is synthesized. Eventually, in the reciting the review stage, the dispersion mechanisms are prescribed, and the review report is offered. The conditioning concerning the planning and conducting of our methodical review are described in the following subsections. The reporting of the review stage is offered in Section 4.

i. Research Question

We've carried out a methodical literature review utilizing the path alluded to. The thing of our study is to examine the current use of usability evaluation styles in trap evolution from the point of prospect of the following exploration question What usability evaluation styles have been assumed by experimenters to estimate trap vestiges and how were they assumed? The criteria exercised to codify the evaluation styles are offered in Section 3.3. This exploration question will help us to epitomize the current knowledge about trap usability evaluation and to identify hiatuses in current exploration to suggest areas for further disquisition.

The study's population and intervention are as follows
• Population trap usability full exploration papers
• Intervention Usability evaluation styles
• outgrowth No seat on the outgrowth itself
• Experimental project Any project Our review is more restricted than a full methodical review as alluded to since we didn't follow up the sources in papers.
In extension, we didn't carry other sources similar to specialized crashes, working out papers, and PhD theses. This program has been exercised in another methodical review conducted in the trap Engineering field.

ii. Identifying and Selecting Primary Studies

The main sources we exercised to probe for primary inquiries are IEE Explore and ACM digital libraries. In extension, we've comprehended the actions of the following special effects and conferences.

• World Wide trap congregation actions – WWW (2003, 2004, 2007), Usability and availability & trap engineering track.
• transnational congregation on trap Engineering actions – ICWE (2003-2007)
• IEEE Internet Computing Special conclusion on " Usability and the trap "(1 measure published in 2002)
• A book on trap Engineering by Springer (LNCS) published in 2005.
• transnational trap Usability and Availability factory actions – IWWUA (2007)

The hunt lacing outlined for reacquiring inquiries is as follows usability AND trap AND evolution AND(evaluation OR trial OR study OR testing) We experimented with several hunt lacings and this one recaptured the topmost quantum of applicable papers. This hunt lacing was exercised in the IEEExplore and the ACM digital libraries as well as in the other sources that were audited manually. The period examined was the last 10 times, i.e., inquiries published from 1998 to 2008. Concerning the digital libraries, we ensured that our hunt program was applied to depositories, diurnals, and congregation actions.

iii. Inclusion Criteria and Procedures

The researchers conducting the systematic review assessed each identified study to determine whether or not it should be included. Consensus was used to resolve the differences. Included were the studies that satisfied the following criteria:
• Papers outlining a usability evaluation method or methods for developing Web applications. The only studies that made use of a "formal" approach (such as cognitive walkthrough or heuristic evaluation) were chosen.
• Comprehensive research reports.
Papers of the following categories were not accepted:
• Articles offering guidelines and suggestions for web design.
• Articles outlining methods for combining usability tests.
• Articles that provide metrics for usability.
• Prefaces to books, workshops, and special issues.
• Works not composed in English.

iv. Data Extraction Strategy

The following criteria, which break down into research questions, were used to compare the extracted data:

1. What techniques have researchers used to assess Web artifacts using usability evaluation methods (UEMs)?

i. Is it a brand-new assessment technique or an established technique from the HCI industry? (Recent, Current)
ii. What kind of usability assessment technique is being used?
(User testing, inspection technique, and other)
2. What stage does the assessment method get used in? (Specifications, Plan, and Execution)
3. What kind of evaluation is it? (Automated, Manual)
4. Was the process of evaluation reviewed? (In agreement, in disagreement). If so, what kind of assessment was carried out? (Case study, experiment, survey)
5. Was it the goal of the evaluation to offer suggestions for improving the design?

About the first criterion, the paper is classified as new if it presents at least
one evaluation system that's specifically drafted for the Web. Otherwise, it's
classified as being if the paper uses styles from the HCI field. In addition, the evaluation system is classified according to the following types examination system, stoner testing, or other. The paper is classified as an examination system if it reports an evaluation grounded on expert opinion(e.g., heuristic evaluation, guideline reviews, norms examination, cognitive walkthroughs). Otherwise, the paper is classified as stoner testing if it reports an evaluation that involves the stoner's
participation. Similar evaluations generally concentrate on lower-position cognitive or perceptual tasks. In this order, we also consider the several protocols that live to conduct stoner testing (e.g., allowing audibility, and question-asking). Eventually, the paper is classified as others if it reports the use of other styles (e.g., focus group, web operation analysis). About the alternate criterion (the phase in which the evaluation is conducted), each paper is classified into one or further ISO/ IEC 12207 high-position processes Conditions, Design, and Software Construction (perpetration). The paper is classified at the phase of the condition if the vestiges used as input for the evaluation include high-position specifications of the Web operation(e.g., task models, uses cases, scripts). The paper is classified at the design phase if the evaluation is conducted on the intermediate vestiges of the Web operation (e.g., nautical models, abstract stoner interface models, dialog models). Eventually, the paper is classified at the perpetration phase if the evaluation is conducted in the Web operation. About the third (the type of evaluation conducted), the paper is classified as homemade if it presents a usability evaluation that's manually performed. Otherwise, it is classified as automated. The fourth criterion is related to the evaluation of the usability evaluation styles. Depending on the purpose of the evaluation and the conditions for empirical disquisition, three different types of strategies can be carried out check, case study, and trial. A check is a disquisition performed in retrospection when the system has been in use for a certain period.

A case study is an experimental study and data is collected for a specific purpose throughout the study. A trial is a formal, rigorous, and controlled disquisition. Trials give a high position of control and are useful for comparing usability evaluation styles in a further rigorous way. For evaluations of this type, statistical styles are applied to determine which system is better. Eventually, the fifth criterion is to determine whether or not the evaluation system provides feedback to the developer. The evaluation system is classified as No if it's aimed at only reporting usability problems. The system is classified as yea if it also provides recommendations on how the problems can be fixed

v. Conducting the review

The hunt to identify primary studies in the IEEExplore and ACM digital libraries was conducted on the 22nd of March 2008. The operation of the review protocol yielded the ensuing results

• The bibliographic database hunt linked 338 potentially applicable publications 181 from the IEEExplore and 157 from the ACM digital library). After applying
the rejection criteria proved in Section 3.3, 37 publications were eventually named(11 from IEEExplore and 26 from ACM digital library).

• The homemade bibliographic review of the other sources linked another 72 potentially applicable publications. After applying the rejection criteria, the
following publications were eventually named 14 papers (3 from WWW, 3 from ICWE, 3 from the IEEE Internet Computing special issue, 4 from IWWUA, and a chapter from the book).

Thus, an aggregate of 51 exploration papers were named by our addition criteria. Some studies have been published in further than one journal/ conference. In this case, we named only the most complete interpretation of the study. Other studies appeared in
further than one source. These publications were taken into account only formerly. The search results revealed that exploration papers about Web usability are published in several conference journals from different fields, similar to Computer Interaction (HCI), Web Engineering(WE), and other affiliated fields.

IV. RESULT

The results of our study are presented in Table 1. They've been organized by selection criteria and publication source. The list of papers containing all the data uprooted from the studies wasn't included in this paper due to space restrictions. These results indicate that 45 of the papers reviewed presented new evaluation styles specifically designed for the Web (seeFig.1(a)). For this case, Blackmon et al. proposed the cognitive walkthrough for the web (CWW) system. When compared to the traditional system, this system was set up to be superior for assessing how well websites support stoner navigation and information hunt tasks. In another study, Bolchini and Garzotto proposed a usability examination system for Web operations called MiLE. The system was estimated through two studies that measured the effectiveness, performance, and perceived difficulty of learning the system. The remaining 55 of the studies reported the use of evaluation styles (e.g., cognitive walkthrough, heuristic evaluation, stoner testing).

Table No 1. Systematic review Result

Selection criteria		IEEE	ACM	WWW	ICWE	IE3IC	Book	IWWUA	T
Usability Evaluation Method	New	4	9	2	3	3	0	2	2
	Existing	7	17	1	0	0	1	2	2
Type of Usability Evaluation Method	Inspection method	4	5	0	1	1	1	1	1
	User testing	7	17	1	0	0	1	0	2
	Other	4	11	2	2	2	1	3	2
Web development phase	Requirements	1	1	0	0	0	0	1	3
	Design	4	4	0	1	3	1	3	1
	Implementation	7	25	3	3	1	1	1	4
Type of evaluation	Manual	9	19	0	1	1	1	4	3
	Automated	2	7	3	2	2	0	0	1
Validation?	Survey	0	3	0	0	0	0	0	3
	Case study	1	3	2	1	0	1	3	1
	Experiment	2	10	0	0	0	0	1	1
	No	8	10	1	2	3	0	0	2
Feedback to design?	Yes	4	6	0	0	2	0	3	1
	No	7	20	3	3	1	1	1	3

IEEE – IEEExplore electronic database
ACM – ACM digital library
WWW – World-Wide Web conference from 2003 to 2007
ICWE – International Conference on Web Engineering from 2003 to 2007
IE3IC – IEEE Internet Computing Special Issu Usability and the Web
Book – A book on Web Engineering by Springe
IWWUA – International Workshop on Usability and Accessibility 2007

The results also revealed that the most constantly used type of evaluation system is stoner testing, i.e., 41 of the papers reviewed reported some kind of testing involving druggies(see Fig. 1(b)). This may indicate that utmost evaluations are performed substantially during the late stages of the Web development lifecycle. Examinations account for 20 of the studies, whereas 39 of the studies reported the use of other styles(e.g., paper prototype, remote stoner testing, check). An illustration of the use of examination styles is described in Sutcliffe. The author proposed a set of heuristics for assessing the attractiveness of Web stoner interfaces. The heuristics were tested by assessing three airline websites. The results of the study show that aesthetics may play an important part in original visits but content issues may be dominant for reprise visits. The analysis of the results verified that the evaluations are substantially performed at the perpetration position of the Web operation(see Fig. 1(c)). Around 27 of the studies describe evaluations performed using the Web operation's intermediate vestiges(e.g., abstract stoner interface, nautical model). Only 5 of the evaluations were performed at the conditions specification position (e.g., laboratory stoner testing of paper mock-up prototypes). Thus, there's a need for usability evaluation styles that can be used at the early stages of Web development. About the type of evaluation, 69 of the studies performed the evaluations manually(see Fig. 1(d)). Around 31 of the studies reported the actuality of some kind of automated tool to support the proposed system. For this case, Becker and Berkemeyer proposed a fashion to support the development of usable Web operations. The fashion is supported by a GUI-grounded toolset called RAD- T(rapid-fire operation design and testing) that allows early usability testing at the design stage. We also vindicated whether the studies reported some kind of empirical evaluation. The results revealed that 47 of the studies didn't conduct any type of evaluation see Fig. 1(e)). Still, it was surprising to observe that, from the papers that performed evaluations, 25 of them reported on controlled trials. The maturity of these studies was published in HCI conferences and journals; hence, the trial is a common exploration system used in this field. An illustration of this is the study conducted by Hornbæk and Frøkjær, where two psychology-grounded examination ways(cognitive walkthrough(CW) and conceits of mortal thinking MOT)) were compared. The results show that the actors linked 30

further usability problems using MOT. Around 22 of the studies report case studies. For
case, Matera presented a case study in which three styles were applied to the evaluation of a Web operation design examination to examine the hypertext specification, web operation analysis to dissect the stoner geste, heuristic evaluation to dissect the released prototypes, and the final Web operation.

Eventually, 71 of the studies reported only usability problems giving no feedback on the corresponding design vestiges(see Fig. 1(f)). The remaining studies also offered suggestions for design changes grounded on the usability problems detected.
For this case, Hornbæk and Frøkjer reported a trial aimed at comparing the assessment of both usability and mileage of problems and redesign suggestions. The results of the trial showed how redesign proffers were assessed by inventors as being of advanced mileage rather than just problem descriptions. Usability problems were seen more as a help in prioritizing ongoing design opinions. Figure 2 shows the number of named publications on Web usability evaluation styles by time and source. The analysis of the number of exploration studies on Web usability showed that there has been a growth of interest in this content. Outmost of the studies about Web usability were set up at the ACM digital library.

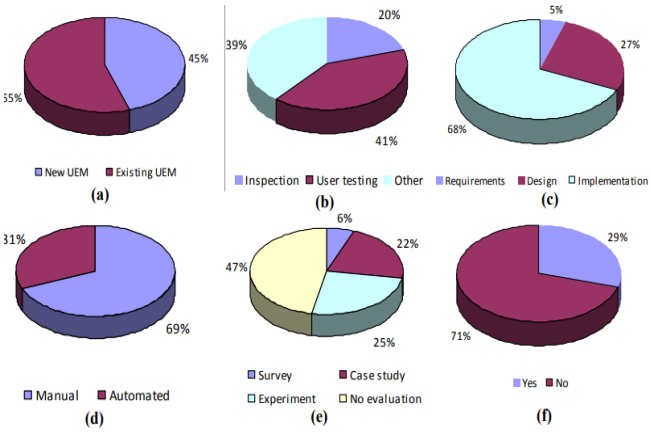

Fig. 1 Percentage of coverage by criteria used for data extraction

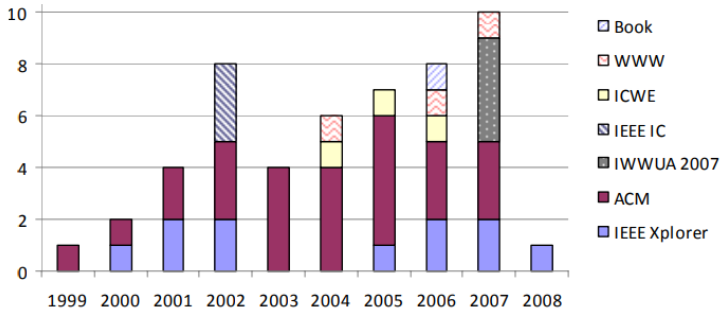

Fig.2 Number of Publications on Web Usability by Year and Source

V. THREATS AND VALIDITY

The main limitations of this study are publication selection bias, trip in data birth, and misclassification. Publication bias refers to the problem that positive results are more likely to be published than negative results. We believe that we have soothed this trouble, at least to some extent, by surveying applicable journal special issues and conference proceedings. Still, we didn't consider slate literature or unpublished results. Regarding publication selection, we chose the sources where papers about Web usability are typically published. Still, we've barred some journals in the Web Engineering field from this methodical review (i.e., Journal of Web Engineering and International Journal of Web Engineering and Technology) since we had no access to these journals. This fact could affect the validity of our results. We tried to palliate the pitfalls of the trip in data birth and misclassification by conducting groups of papers with three pundits.

VI. CONCLUSION

This paper has presented a methodical review of usability evaluation styles for Web development. The results of the review have linked several exploration gaps. In particular, usability evaluations should be performed beforehand in the Web development process and should be done constantly throughout the design cycle, not just when the product has been completed. The maturity of the papers reported on evaluations at the perpetration phase. It also reveals that the evaluations are substantially performed in a single phase of the Web operation development. Usability evaluation at each phase of the Web operation development is critical for ensuring that the product will

be used and be effective for its intended purpose(s). In addition, the maturity of the styles reviewed only allowed the generation of a list of usability problems. New proffers for redesign that address usability problems as an integral part of the evaluation system are demanded. Although our findings may be reflective of the field, further reviews are demanded to confirm the results attained. Unborn work includes the extension of this review by including other sources(e.g., Science Direct and Scopus databases). We also want to dissect more in depth the position of integration of the usability evaluation styles into the different processes of the Web operation lifecycle. Eventually, we plan to collect further information about the empirical substantiation of the effectiveness of usability evaluation styles for the Web.

VII. REFERENCES

[1] Neuwirth C. M., Regli S. H. IEEE Internet Computing Special Issue on Usability and the Web, Vol. 6, No. 2, March/April 2002.
[2] Weske M., Hacid M. S., Godart C. (Eds.): Web Information Systems Engineering - WISE 2007 Workshops Proceedings, Nancy, France, December 3, 2007, LNCS 4832, Springer
[3] Lowe D., Gaedke M. (Eds.): Proc. of the International Conference on Web Engineering 2005, Sydney, Australia, July 27-29, 2005, LNCS 3579, Springer.
[4] ISO – International Standard Organization, ISO/IEC 12207: Standard for Information Technology – Software Lifecycle Processes, 1998.
[5] Lovelle J. M. C., Rodríguez B. M. G., Aguilar L. J., Gayo J. E. L., Ruiz M. P. P (Eds.): Proc. of the Int. Conf. on Web Engineering 2003, Oviedo, Spain, LNCS 2722, Springer.
[6] Baresi L., Fraternali P., Houben G. (Eds.): Proc. of the International Conference on Web Engineering 2007, Como, Italy, July 16-20, 2007, LNCS 4607 Springer 2007.
[7] Becker S. A., Berkemeyer A., Rapid Application Design and Testing of Web Usability. IEEE Multimedia, 9(4): 38-46, October/December 2002.
[8] Fenton, N., and Pfleeger, S. L. Software Metrics: A Rigorous and Practical Approach, Second Edition. International Thomson Computer Press, 1996.
[9] Hornbæk K., Frøkjær E. Two psychology-based usability inspection techniques studied in a diary experiment Proc. of the 3rd Nordic conference on Human-computer interaction (NordCHI'04), Tampere, Finland, pp. 3-12, 2004.