

A competitive analysis of High fidelity prototype with current ‘UNI-WEBMAIL’ System over performance.

Mohammad Izabul Khaled(119013), Asif Iqbal (114313), and Darpan Vats (119188).

Abstract—In this study, we show low complexity wireframe for improved performance and reliable on WEB based Electronic mail system.

Our position is to evaluate the performance comparison of current WEB based Electronic Mailing system and, proposed improved WEB based Electronic Mailing system. Main motivation for this is to achieve – Efficiency with minimalistic error rate and high satisfaction, with the User Interface. The comparison proposed here employs Student T-test with minimum resource contention. WEB based Network elements employ key features of usability-Learnability, Efficiency, memorability, satisfaction and error rates. We found the tests conducted on these usability parameters are, which concludes that the new interface might influence the performance of the primary and secondary user groups. A high-fidelity prototype is presented and the usability evaluation of the working system is compared, with the current WEB E-mail system at Bauhaus University to conclude the performance result.

Index Terms—WEB, software, student T-test, paired t test, E-mail, Mailing, Usability, Interface, Design,, evaluation

1. INTRODUCTION

Interface of an application has the potential to improve the Usability of a software product and the thus, holds the potential to reduce recurring cost on the long terms. In the recent times software stakeholders have been paying attention to the evaluation of software products, to increase the life of a product, with least maintainability. In this era of Internet age, a significant advancement in WEB based applications and portals are recognized. There are several factors concluding the life and usability of the software, amongst which User Interface is one of the major factor, beholding the key role to bridge the gap between the system and the goals user wants to achieve.

Our study apprises the Usability evaluation of two WEB E-Mail services, established on student t-test methods. To simulate the study a high-fidelity prototype has been developed and, tested with the current system, under the natural environment keeping the effects of confounding factors minimalistic. Empirical Study has been evaluated over

the quantitative methods and, the results are abstracted using the graded tools to achieve and concluded with dependent variables.

The paper is organized as follows. In Section 2, authors will start by presenting the formulation of research and, the concluding hypothesis. Section 3 describes the overview of data and the factors contributing to the analysis. In Section 4, discuss the experimental setup and the administration. The evaluation of the proposed study of two systems is outlined in Section 5. Section 6 concludes the outcomes to foster the hypothesis of the test

2. MOTIVATION AND HYPOTHESIS

In this experiment, we wanted to compare the performance analysis to verify the usability of the new develop system in order to know whether of its performance is better or not. Hence, we tried to manage our subjects size relatively large (20-30) for a valid result and select based on their experiences of using the existing system as it is known that experience user benefits more in the usability testing of comparing two system(e.g. UI). We set our research question for the experiment: *Does the new developed ‘UNI-WEBMAIL’ system improves performance?*.

2.1. Hypothesis

From the research question, we conclude hypothesis as we are trying to compare the performance of two systems and the hypothesis are:

H0: ‘Both system are same in performance’

H1: ‘New Developed system of ‘UNI-Webmail’ performance is better than the old one’

Afterwards, we prepared our study design based on above hypotheses.

The next section discusses about the experimental set-up, and the analysis method for the aggregated data to draw the concluding factors.

3. DATA COLLECTION

We are going to compare the performance of two system, we make our variables after setting the hypothesis. Since performance is latent variable, it should be operationalized in order to obtain the variables that will be measured [1]. These are the independent and dependent variable given below.

3.1. Variables

Independent	Dependent
Old webmail system/new developed webmail	Time on tasks-completion, Satisfaction

There are many confounding variables/factors affecting the data to be biased will result with an invalid output. These factors can be: experience, learning effect, environmental influence, various display e.g. tab/mobile/monitor and the keyboard layout etc. Controlling those factors are very important to reduce/avoid the influence by making most of the factors constant e.g. experience. In order to reduce the learning effect, we tried to maintain one 1-week gap in between testing two systems for most of the user and most of the studies were conducted afternoon. In addition, We had used the same instrumental setup for all the participants.

3.2. Validity

For internal and external validity of the study, we were very careful as these validities are very important to evaluating the validity of a research studies. Hence, to validate our output internally, we tried our best to control the above-mentioned confounding factors. Plus, we tried to improve external validity by setting experiments in a more natural way besides taking random sampling of the subjects/participants. [2]

3.3 Choosing test

Since the data is interval/ratio type, the common summary measures are mean, median and standard deviation. It is known that when the population means of only the two groups is to be compared, the *t-test* is used. Hence, based on our data, samples and study purpose, in our experiment, we will be conducting within subject design as two systems will be experimented by same subjects. Moreover, within subject is

simple and more data can be gathered in shorter period of time. Therefore, paired sample t-test would be used for comparing the mean performance. We have planned to test with a relatively large groups of students ($n > 20$) to ensure the data to be normally distributed (if not perfect, at least approximately) and for the meaningful result.

There will be a question after each system testing as we are going to find out the satisfaction level. [1]

4. EXPERIMENT PREPARATION

As part of the experiment, we came up for the experiment preparation based on what we will be able to collect the data directly from the subjects. We thought about two steps for this. In the first step, we have developed our new system/apps which will be used to accomplish some specific task from where we collect data. There is some specific task we set for the participants. Secondly, there will be some questionnaire in order to get feedback about the systems from the subjects.

4.1. High-fidelity prototype

We started developing our prototype to give the user a real feel of the system. It is important to mentioned that we have not developed the whole system for the shortage of time. However, we tried to figure out the important functions in the existing system that are frequently used by the subjects and develop that particular functionality only in the new system. For the development tool, we have chosen Just In Mind 1 month free version, a tool for high-fidelity prototypes of the web, mobile apps and website wireframes [6]. It offers all the necessary things required prototyping which can be simulated on the actual devices as a real product.

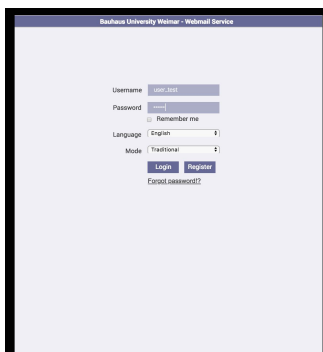


Fig 1: Login attempt

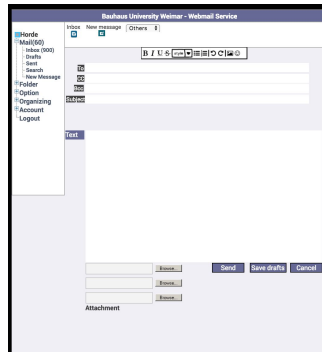


Fig 2: Sending an email

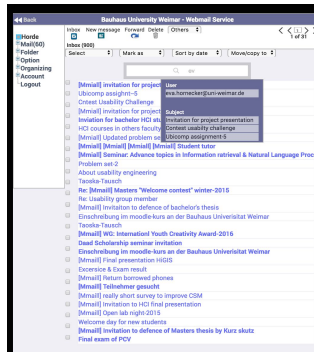


Fig 3: Searching

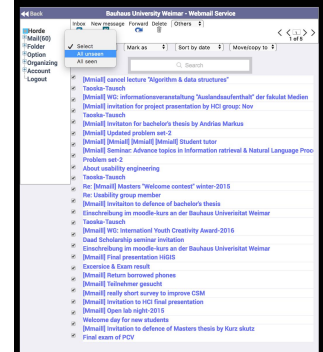


Fig 4: Viewing all unseen

4.2. Questionnaire

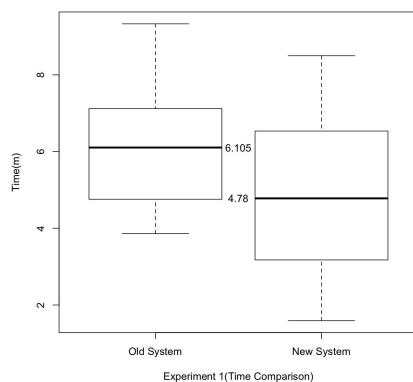
When subjects finish the given task, there will be a question after each system testing as we are going to find out the satisfaction level. This data gives us the perception of the subjects against the systems. For that reason, we used our online questionnaire where user can give their feedback. [2]

4.3 Experiment Conduction

We took almost two weeks for conducting the experiment as we tried to blank some space in between existing and new system in order to reduce learning some learning effects. Almost all the subject has been chosen randomly and they all are Masters student of Bauhaus University Weimar who has at least six months of experience of using the uni-webmail. We had scheduled their free and invite them for the study. As we were having lack of time and students were leaving city for vacation, we sometimes met them at their proposed location. We did not tell anything of our purpose of study by not revealing the research question. We have started the experiment by giving a short explanation for the participants to make sure that they clearly understand the to-do tasks. There were five different tasks we have set for them based on what the data has been collected and these tasks are given below:

- Login [Fig. 1]
- Sending an email to after attaching a file [Fig. 2]
- Searching, opening a particular email [Fig. 3]
- Viewing all unseen/seen emails, opening one among and deleting [Fig. 4]
- Logout

Stopwatch for measuring the time taking to complete all tasks until they logout. For the error, we just count how many times they encounter different problems and asks for help of the facilitator/experimenter. As our focus is time measurement, we made the time limit open until participants are not accomplished with the tasks leaving error rate skipped.



5. EVALUATION

5.1. Participants

Total twenty (20) participants has participated in the study where 13 were male and 7 female current masters students. From the survey, we found that the the age group is in between 23 - 30 and they all are using Bauhaus Uni-webmail for 1 year.

5.2. Test Experiment

In this experiment, paired t test was performed using task completion time for two different system. In addition, we assumed that one system performance would be better than other one, so we selected one-sided paired t-test [4]. There were three tasks namely sending email, searching and viewing emails. We took the aggregated time from all tasks to compare between old and new system. From taking a look in the box plots we can easily observe that there is a clear difference between two tested systems [Fig. 5]. As the median for the old system is 6.105 whereas median for the new system is 4.78 [3].

For the evaluation of paired t test was performed because users were required to complete all the tasks in both systems and it was one tailed as we have predicted the mean performance. We calculated the confidence interval by 95% level for the mean difference to tell us within what limits the true difference is likely to lie. T table has been used for finding the critical value based on dof and significance level.

5.3. Test Questionnaire

Our second data source was test questionnaire filled up by participants after completing the experiment. We used the average of three questions. The interesting fact is the subjective judgement for the different systems. From the box plot we found that there is an almost significant difference between user feedbacks for two different systems [Fig. 8]. From the feedback data 85% of users preferred the new mail system, whereas 10% users found no difference between both systems and only 5% users said they liked the old mail system.

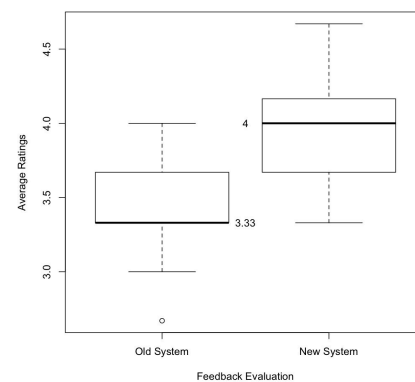


Fig 5 : Task completion time

5.4. Interpretation

The study of Experiment test conducted reveals that the newly developed system has significant performance improvement and likelihood. The finding was derived from time differences of users in old and new system and average ratings from the questionnaire data. Also, we can see in the user feedback rating almost 85% user preferred the new system.

Task Experiment	Mean	Median	p-value(T Test)
old system	6.1725	6.105	0.004527
new system	4.7840	4.78	

Table 1: Mean and Median of the task experiment, as well as the results of the paired t test. Data is normally distributed and shows the significance

Referring to the research question and hypotheses, in 95% confidence interval or corresponding .05 significance level, we found that there is no evidence that null hypothesis H_0 is valid (or there is significance evidence against the null hypothesis as the p value is lesser than significance level of 0.05), so it is rejected. Therefore, we can accept the alternative hypothesis H_1 since the difference between the mean is significant means the 'New Developed system of 'UNI-Webmail' performance is better than the old one'.

But as we know in general hypothesis regarding performance cannot be completely accepted or rejected.

6. CONCLUSION

In this paper, we reported concluding results of High-fidelity WEB based E-Mail prototype and the current

Fig 6 : Feedback Evaluation

WEB based E-Mail system at Bauhaus University. To attain the results the experiment was setup in the natural environment of the test subject, other confounding factors were also taken care of during the experiments. The evaluation of quantitative data

over the normalized distribution found high-fidelity prototype **very very** significantly *(Percentage) usable then the current WEB E-Mail system. Hence, the proposed H-F Prototype is concluded to be more usable on the usability parameter and holds the potential to improve the life to WEB E-Mail system implemented.

REFERENCES

1. <https://learnche.org/pid/latent-variable-modelling/what-is-a-latent-variable>
2. <https://www.simplypsychology.org/validity.html>
3. <https://www.surveymonkey.de/r/Y72MHSV>
4. <https://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm>
5. <http://www.statisticssolutions.com/should-you-use-a-one-tailed-test-or-a-two-tailed-test-for-your-data-analysis/>
6. <https://www.justinmind.com/usernote/tests/32419965/32445631/32445635/index.html>