SIGDOC: G: Conditional Usability Testing for UX Optimization

Nupoor Ranade
North Carolina State University
Raleigh, US
nupoor.ranade@ncsu.edu

ABSTRACT
User experience research has become more important than ever. Widely adopted for large projects, product developers resist usability testing for smaller fixes due to the time and costs they need. What can we do to make tests easier and cheaper? This paper describes a novel 'conditional usability testing' approach to optimize the usability testing process to save resources. This approach is an extension of conventional usability practices. It has improved the efficiency of the usability test described in this study by 170%. By conducting deeper task-analysis and user-analysis, this method provides tools to conduct more tests with less resources. It also helps us extract real user data useful for future testing and persona designing requirements.

Author Keywords
User experience, conditional design, HCI, UX optimization, algorithms

CCS Concepts
•**Human-centered computing** → **Usability testing;**

1 INTRODUCTION

Time is an essential commodity for every practitioner. Time, cost and quality (The Iron Triangle) have been inextricably linked with measuring the success of project management since the 1950s. Time, overarches the chance of other criteria, post implementation from being included [2] in the project management triangle. New parameters have been developed over time for project management, however time remains to be a significant entity directly impacting the quality of projects. There are several examples on how practitioners are engaged in reducing time required for processes, for example, code optimization in software development to save time; getting search results in the least amount of time. Google's Search even displays the amount of time required to produce millions of results after crawling websites on the entire Internet. A user may not visit all the resultant web pages, but the time depiction is a rhetorical move that helps Google build credibility and promise of being the fastest search algorithm on the Internet thus enhancing users' experience [15].

Time is a crucial resource for user experience research as well. User experience (UX) researchers and design professionals leverage usability testing to get user feedback on their product's user interface. Both researchers and users have to provide sufficient time for the tests to happen. Despite their expertise with user research, why can't researchers simply design intuitive and effective products to save time? Norman and Nielsen explain that studying users is more complicated owing to differences in user characteristics and the nature of tasks [13, 12]. The objective of usability tests is to observe users perform such tasks to reveal areas of confusion and uncover opportunities to improve the overall user experience [16, 3, 12]. The basic goals of usability testing vary by study, but they usually include [3, 19]: Identifying problems in the design of the product or service, uncovering opportunities to improve and learning about the target user's behavior and preferences.

2 PROBLEM AND MOTIVATION

To attain usability testing goals, usability experts recommend iterative usability testing [6, 12], a well-known technique to repeat usability evaluations by performing design and testing repeatedly until you get a successful version. The purpose of iterative testing is to find as many usability problems in a user interface design as possible to correct design problems in future versions. Although usability testing is cheaper than before [12], there are costs associated with each iteration. How can usability evaluations be made cheaper? One solution is to reduce the number of evaluations directly impacting the number of iterations which could have potentially identified design issues. Another solution is to reduce the number of users in each iteration. However, that impacts the inclusivity of the participants' sample available for testing. These challenges make the test less effective [16]. To address these concerns, this study attempts to demonstrate an alternative testing approach based on the conditional design method [10]. The empirical research that led to this study was implemented at SAS Institute, a data analytics company headquartered in Cary, North Carolina. The conditional usability testing method described in this study helps identify participants' characteristics that can not only enable to save testing time but also find ways to make the product design more inclusive. The goal of this method is to derive the most valuable results from the least

ACM ISBN X-XXXXX-XX-X/XX/XX.
DOI: http://dx.doi.org/10.1145/0000000.0000000

number of iterations. Specifically, this study examines the following research questions:

- How to design conditional usability tests based on user characteristics and testing tasks?
- How much time can be saved using conditional usability tests?

3 BACKGROUND AND RELATED WORK

Researchers have strived to improve usability testing processes [19]. But most of them have focused toward getting more results. This study focuses on improving the process of conducting usability tests to save the resources spent on them. Usability testing is performed in multiple ways, but the most common elements involved are the facilitator, the tasks, and the participant [3, 13]. Each of these components need to be studied closely to make a change in the process of conducting usability tests.

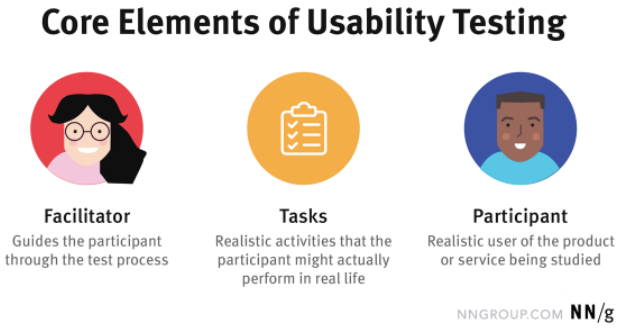

Figure 1. A usability-testing session involves a participant and a facilitator who gives tasks to the participant and observes the participant's behavior [3, 13]

Previous researchers have examined the role of users in two ways

- User personas used to develop the product as well as a UX study [18]
- Number of users required for a study [12, 7].

Personas are fictional representations and generalizations of a cluster of the target users who exhibit similar attitudes, goals, and behaviors in relation to your product [4]. Although personas are useful in educating stakeholders about the product goal, purpose and design, they may not be useful in the participant recruitment process for usability studies. First, getting participants who fit the persona's template is difficult. Second, personas fail because they are a broad depiction of users [12] when in reality every user might be unique. While users may be different, Norman and Launderer's studies have proved that the best results come from performing usability tests on no more than 5 users and running as many small tests as can be afforded. As the number of users goes beyond 5, we start noticing similar or repetitive results on some tasks which do not contribute in any way to make design decisions. However, if researchers have a larger participant body at their disposal, Norman and Launderer suggest breaking them up into smaller groups and performing iterative or comparative testing [7, 11]. To make use of these suggestions, the model of usability tests needs to be reconsidered.

Iterative testing (as mentioned earlier) is the process of conducting one test after another while revising each version based on the usability findings of the previous. New test cases (tasks) are designed for each iteration. Researchers have observed usability to be improved by at least 20% with each iteration [13, 14]. Multiple iterations can make testing expensive and time consuming, so researchers have suggested other approaches like parallel and comparative testing. Although these approaches reduce the number of tests, they require usability researchers to have alternative test cases that can be used for testing. For example, if a button on a website is being tested, parallel and comparative tests need researchers to have multiple variations of the button's designs handy so that user's opinions for all of them can be determined at the same time. Developing these designs requires time and human resources which can lead to delays in scheduling usability tests. Norman and Nielsen suggest that the 3 methods for increasing UX quality by exploring and testing diverse design ideas work better when used together [18]. Although combining them might improve results, they do not help in reducing costs as much as the number of tasks and participants remain the same. We need to find alternative approaches that pay more attention to users and tasks, two of the three main elements of usability testing.

This study combines user-analysis and task-analysis to design a new alternative approach of conducting usability testing. In this approach, a conditional test algorithm is used to help facilitators make decisions on-the-fly and use those decisions as a guide for picking next tasks during the testing process. The conditional test algorithm is based on the conditional design method [8] formulated by the graphic designers Luca Maurer, Jonathan Puckey, Roel Wouters and the artist Edo Paulus, in which conditions are adapted based on participant interactions. In conditional designs, emphasis is on the process instead of the product. Logic and intelligible rules are used to design the conditions through which all the processes take place [8]. In usability testing, user characteristics, subjectivity of test participants [17], test tasks and testing processes can be used to build a conditional design testing approach. The process of creating such a test and results of implementation have been shared in the following sections.

4 METHOD

The conditional usability method, designed by refining the level of user participation [7] in the testing process, is a three-step process that involves planning, implementation and reporting [1]. This approach was implemented and tested at the SAS Institute. Tests were conducted on the features of SAS's product documentation website called Help Center. The following steps were followed to conduct usability tests using the conditional test algorithm.

Planning

A single test iteration was designed to test several features of the Help Center. The test was scheduled for 5 minutes (excluding pre-test and post-test tasks). For this study, 10

Table 1. Number of tasks completed (Test A)

Participant	Tasks completed	Total tasks	Total test time (minutes)
A	4	4	2
B	3	4	5
C	4	4	4
D	2	4	5
E	4	4	5

participants were recruited. User-analysis and task-analysis were performed in the following ways:

- Task Analysis
  Tasks were defined by the researcher based on her own experience with usability testing the same tasks with previous participants.
  - Difficulty levels: Features known to site users were classified as **easy**, common design features were classified as **medium** and new features that require training were classified as **hard**.
  - Testing priority: High priority testing tasks were classified as **primary** and low priority testing tasks were classified as **secondary**.
- User Analysis
  Users were categorized as well.
  - Tasks performed: Classification on participants' abilities to perform tasks as **beginners**, **medium** and **advanced**
  - Personas for future use: Based on Albers' model of multidimensional user analysis [1], personas were defined into the following categories:
    Dimension 1 **Knowledge** (Corresponding to background knowledge or education if revealed), **Details provided** (Level of details provided during the think-aloud process), **Experience using product** (Years of experience using the product if revealed)
    Dimension 2 **High , Medium and Low**
    After the testing, users information was saved in form of their persona for future use. The persona was a combination of Dimension 1 and Dimension 2. For example, Peter, High knowledge level, low details

Implementation

Lewis notes that for iterative testing a stopping role should be defined based on the availability of time and economic resources [4]. The 5 minutes of scheduled test time was used as the stopping rule [7, 11, 12] and an algorithm for the process was designed. The same algorithm and list of tasks were used to develop the moderator script for the test. Testing was based on a pre-existing set of rules defined by other scholars [11, 7]. All the usability tests were facilitated by the author. All participants evaluated the same web site. All used remote testing and the think-aloud method. All participants received the same instructions prior to the evaluation.
Out of the total 20 tasks, 4 were primary and 16 were secondary. A large number of secondary tasks were used to ensure maximum utilization of the scheduled test time. Out of 16 secondary tasks, 3 were hard, 5 were medium, 8 were easy. Note that the number of tasks picked were inversely proportional to the time they take for execution. For example, easy tasks take less time, therefore more of them were included in the secondary tasks list.

Algorithm 1: Time-based conditional usability testing

```
Result: Total number of tasks = Ttask
Total time = Ttot
Number of primary and secondary tasks = np and ns
Time for each primary and secondary tasks = tp[np] and
  ts[ns]
Time counter = Tcount, Task counter = Ttask
Primary task = P[np], Secondary Task = S[ns]
cut-off time for primary task = [hard, medium, easy]
// Perform primary tasks
for i<np; i++ do
    tp[i] = perform P[i] // performs task and returns time
      required
    Tcount = Tcount + tp[i] // Time required gets added
    Ttask ++
end
if Tcount > Ttot then
    break;   //exit if primary tasks use up test time
else
    // Perform secondary tasks
    while Tcount < Tot do
        if tp[average(task type)]< time sec task[task
          type]) AND (number of secondary task of that
          type>0) then
            ts[hard task] = perform S[hard task];
            Tcount = Tcount + ts[hard task];
        end
    end
    Ttask++
end
```

Testing

During the test, time required for each type of task and total time were recorded. Two tests were conducted with 5 participants in each [5, 13, 7].

- Test A
  A standard (conventional) [3] usability test was conducted with only 4 primary tasks. The time was set to 5 minutes.
- Test B
  The algorithm was followed for conducting the test and data was recorded in a spreadsheet. The persona of participants was recorded and stored in the same spreadsheet.

Table 2. Number of tasks completed (Test B)

Participant	Primary tasks completed	Secondary Tasks			Total tasks	Total test time (minutes)
		Hard	Medium	Easy		
A	4	2			6	5
B	4			4	8	5
C	4		5		9	5
D	4				4	14
E	4			3	7	5

At the end of the tests, results of standard testing (Test A) and conditional usability testing (Test B) were compared. Results were used to calculate the efficiency of the conditional testing method.

Reporting

While reporting results, the design related problems discovered in the usability testing process were documented. This information included the time required to perform each task, number of tasks achieved by each user and users' recommendations for improving the Help center design.

5 RESULTS

Usability testing was conducted by allocating time constraints to ensure a high-rate usability testing system. First the test was conducted for primary tasks.

Test A

In all, 17 tasks were completed using the standard usability testing approach.Some tasks took lesser amount of time (see Table 1). For example, participant A completed all the tasks in 2 minutes, but the remaining time was wasted as the stopping rule was only imposed on the time and not on the tasks.

Test B

Since the total scheduled time was 5 minutes, the remaining time was used for testing secondary tasks. The time taken for each primary task was used to determine the difficulty of the secondary task. For example, if the average time for completing hard primary tasks was less than 25 seconds, the participant was tasked with only hard secondary tasks. Similarly, the cut-offs for medium and easy primary tasks were 10 seconds and 15 seconds respectively (see Table 2). In all, 34 tasks were completed using the conditional usability test approach.

6 UNIQUENESS OF APPROACH

It is clear that more number of tasks were completed using the conditional usability approach. There are several other advantages of using the conditional design approach as well. A design problem can be quickly determined without wasting time prompting the participants for directions. Participant characteristics can also be determined simultaneously. More tasks can be achieved in the same amount of time. For example, in this scenario, all tasks take up to 5 minutes. In the same amount of time, a participant's behavior can be tested for varying number of tasks. The efficiency of the design can be evaluated. Efficiency is calculated as the ratio of achieved work to the expected work. Based on Test B results, the **efficiency achieved is 170%**.

$$Efficiency = \frac{Workdone}{WorkExpected} * 100 \quad (1)$$

Iterations

Although the efficiency is about 170%, this number is not completely accurate because standard usability testing approaches assume that the testing is iterative. We can assume that the secondary tasks were meant for the second iteration of testing. But since we were able to obtain results for the secondary test cases for 4 out of 5 participants in this iteration itself, for the second iteration we only need to conduct the test with one participant. This means that the efficiency achieved is actually much more. The model for conditional usability testing can thus be described as a combination for elaborate evaluations [12] and iterative testing [6]. The difference between this approach and standard testing is the need of multiple iterations with all participants.

Real data for future studies

Research has shown the significance of having participant's real data to amplify usability test results [20, 9]. It is known to be a way to reduce the artificial aspect of usability studies and to make their results more valid. Sometimes, participants are screened before recruiting them for usability tests. However, screening tasks can vary from the actual tasks of the test. This results in false positives. Incorporating a process to get users' real data requires additional effort [20]. The conditional usability testing method proposed in this study will help incorporate this additional step easily into a usability testing process. The user analysis approach suggested in this study can be expanded to include more dimensions and more characteristics of users. Having real data can save time and costs even further.

Participant Recruitment

The conditional usability test approach is based on an algorithm which evaluates user characteristics on-the-fly when the test is being conducted. As mentioned earlier, the conditional design approach reduces the number of future iterations. For the case implemented in this study, the second iteration had to be conducted for only one participant. At times, getting the same participant for the next test again may not be feasible for at least two reasons: first, the willingness of the participant to contribute more time; second, when the participant is exposed to the system changes, it is likely that they will perform tasks during the next iteration with a preconceived notion of the expected results. To prevent that, it is best to recruit a new

participant. Although it is not advisable to recruit the same participant, we can benefit from recruiting a participant with similar characteristics. This is where the real data will again come handy. For example, if the participant in the previous study was a beginner, we may be able to hire a participant who falls in the same dimensions as defined by the persona retrieved by previous iterations. The conditional approach helps design personas which can be used for recruiting test participants for subsequent iterations.

7 CONCLUSION

The optimized approach for usability testing is presented in this paper using three steps planning, implementation and testing. Implications of conditional usability tests apply to test participants (less time, more questions, elaborate test results), test iterations (recruiting and testing to gauge detailed information) and the costs (optimization costs calculated) of conducting usability tests. Currently the limitation of this project is the perceived subjectivity [17] used for both – task-analysis and user-analysis. By incorporating more facilitators in future tests, the subjectivity can be minimized. Future work will focus on developing an application that will help ease the reporting process for the facilitator. Instead of manually reporting test results and calculate time, an automated system would be more convenient. More research will also be useful in determining the impacts of subjectivities of participants and facilitators, both.

REFERENCES

[1] Michael J. Albers. 2003. Multidimensional Audience Analysis for Dynamic Information. *Journal of Technical Writing and Communication* 33, 3 (Jul 2003), 263–279. DOI:http://dx.doi.org/10.2190/6KJN-95QV-JMD3-E5EE

[2] Roger Atkinson. 1999. Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. *International Journal of Project Management* 17, 6 (Dec 1999), 337–342. DOI: http://dx.doi.org/10.1016/S0263-7863(98)00069-6

[3] Carol M. Barnum. 2010. *Usability Testing Essentials: Ready, Set...Test!* (1st ed.). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[4] Mary B Coney and Michae L Steehouder. Guidelines for Designing and Evaluating Personas Online. (????), 14.

[5] Joseph S. Dumas, Joseph S. Dumas, and Janice Redish. 1999. *A Practical Guide to Usability Testing*. Intellect Books.

[6] James R. Lewis. 2014. Usability: Lessons Learned ... and Yet to Be Learned. *International Journal of Human–Computer Interaction* 30, 9 (Sep 2014), 663–684. DOI: http://dx.doi.org/10.1080/10447318.2014.930311

[7] Gitte Lindgaard and Jarinee Chattratichart. 2007. Usability testing: what have we overlooked?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*. ACM Press, 1415–1424. DOI: http://dx.doi.org/10.1145/1240624.1240839

[8] Jonathan Puckey Roel Wouters Luna Maurer, Edo Paulus. 2013. Conditional Design - Conditional Design. (2013). https: //www.conditionaldesign.org/manifesto/index.html

[9] Andrew Mara and Miriam Mara. 2015. Capturing social value in UX projects. In *Proceedings of the 33rd Annual International Conference on the Design of Communication - SIGDOC '15*. ACM Press, 1–6. DOI: http://dx.doi.org/10.1145/2775441.2775479

[10] Luna Maurer, Edo Paulus, Jonathan Puckey, and Roel Wouters. 2013. *Conditional design workbook*.

[11] Rolf Molich. 2010. A Critique of "How To Specify the Participant Group Size for Usability Studies: A Practitioner's Guide" by Macefield. 5, 3 (2010), 5.

[12] Jakob Nielsen and Thomas K. Landauer. 1993. A Mathematical Model of the Finding of Usability Problems. In *Proceedings of the INTERACT 93 and CHI 93 Conference on Human Factors in Computing Systems (CHI 93)*. ACM, 206–213. DOI: http://dx.doi.org/10.1145/169059.169166 event-place: Amsterdam, The Netherlands.

[13] Donald A. Norman. 2005. Human-centered design considered harmful. (Jul 2005). http://doi.org/10.1145/1070960.1070976

[14] Liza Potts and Michael J. Salvo. 2017. *Rhetoric and Experience Architecture*. Parlor Press.

[15] Emma Rose and Josh Tenenberg. 2016. Arguing about design: A taxonomy of rhetorical strategies deployed by user experience practitioners. In *Proceedings of the 34th ACM International Conference on the Design of Communication - SIGDOC '16*. ACM Press, 1–10. DOI: http://dx.doi.org/10.1145/2987592.2987608

[16] Jared M. Spool, Carolyn Snyder, and Mavis Robinson. 1996. Smarter usability testing: practical techniques for developing products. In *Conference companion on Human factors in computing systems common ground - CHI '96*. ACM Press, 365–366. DOI: http://dx.doi.org/10.1145/257089.257375

[17] S. H. Stephens. 2017. Designer perceptions of user agency during the development of environmental risk visualization tools. In *Proceedings of the 35th ACM International Conference on the Design of Communication - SIGDOC '17*. ACM Press, 1–7. DOI: http://dx.doi.org/10.1145/3121113.3121223

[18] W. Craig Tomlin. 2018. *UX Optimization Overview*. Apress, 1–10. DOI: http://dx.doi.org/10.1007/978-1-4842-3867-7_1

[19] Ashley Williams. 2019. User-centered design, activity-centered design, and goal-directed design: a review of three methods for designing web applications. (2019), 8.

[20] Todd Zazelenchuk, Sara Sazegari, Kari Sortland, Mark Keavney, and Alex Genov. 2008. Using Participants' Real Data in Usability Testing: Lessons Learned. *Case Studies* (2008), 8.