Cultural Usability: The Effects of Culture on Usability Testing

Qingxin Shi

Department of Informatics, Copenhagen Business School，Denmark
qs.inf@cbs.dk

Abstract. Culture has already played an important role in the global market. It not only affects products, but also impacts on usability evaluation methods. This project aims to examine in the established thinking aloud usability evaluation method (TA UEM), how does the evaluator build a supportive relationship and communicate effectively with the user in order to find relevant usability problems in culturally localized applications. It includes three parts, pilot study, field study and experiments, to get both qualitative data and quantitative data. From this project, we hope to find an effective way to structure our TA UEM methodology to capture or be sensitive towards the mental models and ways of thinking in different cultural groups.

Keywords: Thinking Aloud Usability Testing, Culture, Localization, Evaluator Effect

1 Basic information of the PhD research

Dissertation advisor: Torkil Clemmensen, tc.inf@cbs.dk; Kan Zhang, zhangk@psych.ac.cn
The research area or sub-area of my work: The research area of my work is usability, and the sub-area is cross-cultural usability evaluation method.
Universal resource locator (URL): www.cbs.dk/staff/shi and http://culturalusability.cbs.dk/
Brief description of research topic: It's to investigate the cultural effects on thinking aloud usability testing, focusing on relationship and communication between evaluators and users from same and different cultures.

2 Introduction

With the advent of globalization and IT revolution, we can no longer overlook the aspect of culture in the design of user interfaces and products. In order to capture global markets, the products and software must be tested in target cultures to make sure that they are acceptable and suitable for people's cultural characteristics. But some previous studies have found that culture not only influences the products or

interface design, but also the design methods used in building interfaces [1]. The PhD project will examine the cultural effects on thinking aloud usability testing.

Thinking aloud usability testing method has been extensively applied in industry to evaluate a system's prototypes of different levels of fidelity [2]. The primary goal of a usability test is to find a list of usability problems from evaluators' observations and analysis of users' verbal and non-verbal behavior; thus, the relationship between the evaluator and user is very important for finding accurate usability problems. Tamler [3] suggested establishing a trusting and supportive relationship in order to make the users honestly disclose their thoughts and feelings.

During usability testing, representative users are required to complete pre-established tasks by using the system. This measurement is largely related to specific users and specific tasks. However, people differ across regional, linguistic and country boundaries; therefore, if the evaluator and user have different cultural backgrounds, they may be strongly influenced by their local cultural perspective, perception and cognition, so the interaction and communication between them may be different from those who are from the same culture. Since usability testing involves human-human interaction, the evaluator and user's cultural background must be considered, or else there may be a misunderstanding between them. Therefore, how to build an effective relationship in the usability test has become a key issue in cross-cultural usability testing.

Although the thinking aloud usability test is generally thought to be an effective and successful technique [4], practitioners do not conform to the theoretical basis of the thinking aloud method in the industrial area which was described by Ericsson and Simon [5]. Therefore, Boren and Ramey [5] proposed speech communication theory as a theoretical basis for thinking aloud in usability testing, focusing on evaluator-test user communication. In my PhD project, I also want to examine how the two thinking aloud approaches are used in cross cultural usability testing. From our pilot study, we found that it might be better to follow Ericsson and Simon's approach to do the thinking aloud for foreign evaluators, since they may not be certain what the critical issue is that needs to be probed. In order not to influence the users and get more accurate information, the better way is to interrupt them less and avoid false leading. On the other hand, it might be better to follow the communication approach proposed by Boren and Ramey for native evaluators. Although we have such hypotheses, it is not confirmed by previous researches. This research will try to find what is the effective approach applied in the local and foreign culture in order to find the critical and accurate usability problems of the culturally localized application.

2.1   Research Question

In the established thinking aloud usability evaluation method (TA UEM), how does the evaluator build a supportive relationship and communicate effectively with the user in order to find relevant usability problems in culturally localized applications?

The research question is divided into three sub-questions:
1. Do users produce different thinking aloud protocols when they are with a native or foreign evaluator during the usability test?

2. How are the two thinking aloud approaches (Ericsson and Simon's approach and Boren and Ramey's approach) applied in cross-cultural usability testing? What is the effective approach for local evaluator and what is the effective approach for foreign evaluator?

3. How do we structure our TA UEM methodology to capture or be sensitive towards the mental models and ways of thinking in different cultural groups?

3 Culture Theory

The PhD project will be primarily based on Nisbett's culture theory [6, 7]. His theory focuses on the cognition and perception differences; for example, people from western countries and eastern countries will be different in causal attribution, categorization, and attention to the context vs. salient object [8]. This theory is more relevant to usability testing because thinking aloud usability evaluation methodology asks users to work on typical tasks and to verbalize their task performance and thought process [9]. The whole process involves users' cognition and perception characteristics. The results of the usability test, i.e., usability problems, which are found by the evaluators, are also involved in the evaluators' cognition and perception of the whole test process. When cultural differences exist between the evaluator and test user, some usability problems might be masked, instead of being uncovered.

In my study, temporarily, I define culture as an integrated pattern of local practice knowledge, cognition, and behavior, which is both a result and process of interacting with the environment and society. This project involves three countries: Denmark, India and China. Because Denmark is a western country, and it can also represent the type of society for which usability testing methodology was developed. China is a typical eastern country and India is a multicultural and multilingual country.

4 Main Factors in Cross-Cultural Usability Testing

4.1 Evaluator and User's Cultural Background

Cultural background needs to be considered since users from different cultures may not be influenced to the same degree when they are with a foreign evaluator. Northern European culture is a typical task-focus culture [10], which means that users in those countries may not be influenced when the evaluator is from another country since they pay more attention to the task, not the evaluator. While East Asian culture and Indian culture are socio-emotional relational orientation cultures[10], users in these countries may be influenced more when they are with a foreign evaluator, since the users' effort and attention are directed towards the interpersonal climate of the situation. For example, a study done by Vatrapu and Pérez-Quiñones [1] shows that Indian users who were with a foreign evaluator did not talk as freely as those who were with a local evaluator when testing a localized website. But this may not be the

case for Danish users. In our future study, we will use foreign evaluators in India, China and Denmark to see whether the effect degree is the same in different cultures.

4.2   The Application/Software/Interface Being Tested

The requirement of an evaluator's cultural background is also related to the application or product which is tested in the target culture. There are two approaches to design products for international markets: globalization and localization [11]. "Globalization seeks to make products general enough to work everywhere and localization seeks to create custom versions for each locale" [11,p.158]. If testing a localized application which adapts specific cultural elements for a specific target culture, the results of the usability test may be more related to the evaluator and user's cultural background. Since usually the purpose of doing usability testing in different culture is to see whether this application is accepted by the target culture, which is the localization process, in the PhD research, I still use culturally localized application/software.

4.3   Evaluator Effect

The influence of culture on usability testing may also be derived from another factor called Evaluator Effect: the total number of usability problems found will depend upon the knowledge and experience of the evaluator and the number of evaluators [4]. The evaluator effect indicates that even in one culture, evaluators with different experience will find different usability problems. The effect may be much more significant when the evaluators are from two different cultures, since they do not even have the same cultural background. Even though they are both very qualified and professional, their cognitive process and knowledge cannot be the same, which may be a strong impact factor on cross cultural usability testing.

In a cross-cultural usability test, how can we minimize the evaluator effect which is derived from culture? It is very hard to change the foreign evaluator's cognitive process, but it may be much easier to increase his/her knowledge related to the culturally localized application. The foreign evaluator does not need to master all the target cultural knowledge, because it is impossible. But he can get some important information just related to this application. Maybe he/she needs to know the background, using habits and some related culture features of the application in the target culture, which will be very helpful for them to understand and communicate with the users in the usability test.

5   Methodology

In my PhD project, I will do both field study and experiments to collect both quantitative and qualitative data. Before the field study, we did pilot studies of experiments to collect data in lab conditions in Denmark, India and China, in order to get a general idea of the communication and relation models in the three countries.

Then we did field study in order to examine in the industrial area, how the evaluators do the test and the relation and communication between evaluator and test user in the three countries. Next, I will design some experiments to create a comparable condition to explore the effective relation and communication pattern between foreign/local evaluator and test user.

5.1   Field Study

We conduct field observation in order to find the cultural difference in usability testing in Denmark, India and China. We have finished the field study in India and China and also did one study in Denmark. Maybe we will have one or two more field studies in Denmark. In the field studies which we have done, we went to the companies which did the usability test, and observed their tests.

Until now, we have observed one company in Denmark, one in India and five companies in China. In each company, we did field observation with video cameras of thinking aloud usability test sessions and afterwards interviewed the evaluators, test users and the usability department manager to get more deep information. We used three observer/interviewers: a Dane, an Indian and a Chinese. I will analyze the data in my PhD thesis.

5.2   Experiments

Since methods of experimental design and data analysis contribute more general enterprise of science [12], I will do a series of experiments in Denmark, China and India to figure out such issues:

1. Get a general knowledge of thinking aloud usability testing in Denmark, India and China, and a general knowledge of thinking aloud usability testing facilitated by local and foreign evaluator.
2. Examine the two thinking aloud approaches used in cross-cultural usability testing.
3. Investigate whether training could act as the effective way to structure our TA UEM methodology to capture or be sensitive towards the mental models and ways of thinking in different cultural group.

Application: Culturally localized application.

Variables:

1. Culture: Denmark, India and China.
2. The evaluator and user from the same culture or from different culture.
3. The thinking aloud approach (Boren and Ramey or Ericsson and Simon's approach) that the evaluator follows.
4. The related cultural knowledge that the local and foreign evaluator mastered about the application.
5. Foreign evaluators being trained or not with the target cultural knowledge related to the culturally localized application.

Dependent variables: The number and type of usability problems found by evaluators (severe vs. non-severe usability problems, and shared vs. unique usability

problems), suggestions, positive comments, negative comments and culturally related comments which are made by the users.

6   Conclusion

This project aims to investigate the cultural influence on usability testing from an empirical and theoretical viewpoint. Using foreign evaluator may be much more convenient and less costly in some situation. For example, a Chinese company hopes to extend their products in Denmark. It may be more effective to get feedback to the designers and less costly if using their own usability professionals to do the test in Denmark instead of employing a Danish usability professional. Considering the cross-cultural cost-benefit analyses, it is worth to do the research of how to use a foreign evaluator conducts the usability test in the target country. From my PhD research, I hope to design a new and more effective thinking aloud usability test with less cultural bias for the global market. Besides, this study can also contribute to further understanding of how the two thinking aloud approaches applied in cross-cultural usability testing.

References

1. Vatrapu, R. and M.A. Pérez-Quiñones, *Culture and Usability Evaluation: The Effects of Culture in Structured Interviews.* Journal of Usability Studies, 2006, 2006. **1**(4): p. 156-170.
2. Law, E.L.-C. and E.T. Hvanneberg. *Analysis of Combinatorial User Effects in International Usability Tests.* in *CHI 2004.* 2004. Vienna, Austria.
3. Tamler, H., *High-tech versus high-touch: The limits of automation in diagnostic usability testing.* http://www.htamler.com/papers/techtouch/.
4. Clemmensen, T. and S. Goyal, *Cross cultural usability testing.* Working paper, Copenhagen Business School,   Department of Informatics, HCI research group, 2005. **2005-006**: p. 20.
5. Boren, M.T. and J. Ramey, *Thinking aloud: Reconciling theory and practice.* IEEE Transactions on Professional Communication, 2000. **43**(3): p. 261-278.
6. Nisbett, R.E. *Cognition and Perception East and West.* in *28th International Congress of Psychology in Beijing.* 2004.
7. Nisbett, R.E. and A. Norenzayan, *Cultural and Cognition*, in *Stevens' Handbook fo Experimental Psychology, third edition*, D.L. Medin, Editor. 2002.
8. Nisbett, R.E. and T. Masuda, *Cultural and point of view.* PNAS, 2003. **100**(19): p. 11163-11170.
9. Ramey, J., et al. *Does Think Aloud Work? How Do We Know?* in *CHI 2006, April 22-27.* 2006.
10. Sanchez-Burks, J., R.E. Nisbett, and O. Ybarra, *Cultural Styles, Relational Schemas and Prejudice Against Outgroups.* 2000, University of Michigan.
11. Horton, W., *Graphics: The not quite universal language*, in *Usabillity and Internationalization of Information Technology*, N. Aykin, Editor. 2005, Lawrence Erlbaum. p. 157-188.
12. Maxwell, S.E. and H.D. Delaney, *Designing Experiments and Analyzing Data, A Model Comparison Perspective.* Second Edition ed. 2004, Mahwah, New Jersey: Lawrence Erlbaum Associates.