

# Cross-Cultural Applicability of User Evaluation Methods: A Case Study amongst Japanese, North-American, English and Dutch Users

Vanessa Evers

The Institute of Educational Technology

The Open University

Walton Hall, MK7 6AA

+44 7779 128090

v.evers@open.ac.uk

## ABSTRACT

This paper describes the findings for an international user study investigating cultural applicability of user evaluation methods. The case study evaluates cultural differences in understanding of a virtual campus website across four culturally different user groups by using the same methods for each group. Findings suggest that some user evaluation methods are less applicable than others are for a culturally diverse user base.

## Keywords

International usability testing, culture, user research methods

## INTRODUCTION

Interface localisation becomes more common as companies start marketing their software abroad and aim to target a global e-commerce market. Even though localisation is often thought of as an effort mainly in translation, it is becoming clear that language is only part of preparing a product for use in a different cultural locale. Cultural localisation will often need to involve selection of a suitable real world metaphor, culturally sensitive graphic design and attention to aspects such as flow of information, local content and documentation (see [1] and [2]).

Cultural localisation also involves testing interfaces in the target user cultures and conducting requirement analysis research. Usability testing often consists of one-on-one observations in which a user performs one or more critical tasks by using the website or software (see [3] and [4]). The user's behaviour is observed by the facilitator or researchers and can be preceded or followed by a questionnaire or interview. The study described in this paper aims to provide insight into the need for culturally appropriate research methodologies when investigating international user groups.

Proceedings CHI'2002 Minneapolis, USA

## METHODOLOGY

The case study involved 120 secondary school students from England, North America, the Netherlands and Japan who evaluated the website of a virtual campus ([www.directED.edu/core.html](http://www.directED.edu/core.html), see figure 1). The website was created by a North American firm offering distance education. The sample consisted of 30 students from each nationality, an even number of boys and girls, who were all enrolled in private, fee-paying education.



Figure 1. The DirectED virtual campus homepage.

The user evaluation adopted a multi-method approach. This included a pre-questionnaire in which demographic data was collected and subjects were asked questions about their cultural background and computer experience. There was a task observation session where the participants' were asked to find out whether they would be interested in enrolling at DirectED. They were asked to 'think out loud' while looking for information on the site. Finally, there was an interview, which collected feedback on the website evaluation. This was conducted to validate the observations made during the tasks. More detailed

information on the sample, methodology and instruments can be found in [5]. Each of the participants was exposed to the same methods and instruments even though the instruments were translated for the Dutch and Japanese and facilitation was carried out with translators.

## FINDINGS

Table 1 shows which methods posed complications in data collection for each of the groups involved.

Table 1. An 'X' indicates that the data collection method posed problems for the group involved.

Method	US	UK	NL	JP
Questionnaire	X			X
Task observation (think aloud)		X	X	X
Interview	X			

### Questionnaires

The North Americans found some questions problematic. For the question about what languages the participants spoke, some ticked the option for 'English (British)' as well as 'English (American)'. This shows that even though the use of one questionnaire for different language groups makes the data comparable, it can introduce response options, which are not immediately clear. The other difficulty North American participants had was entering their cultural background. Some would state their religion or race ('white' or 'Christian') rather than more commonly used cultural categories such as 'American' or 'Irish'. For the purpose of the overall study (an investigation into cultural value orientations) this question was extremely informative, however when grouping respondents in cultural categories it may be preferable to offer a comprehensive list so that the most appropriate option can be selected.

The Japanese seemed to associate the questionnaire with a test and needed reassurance to start. Some of the questions posed problems, as it was often difficult to find an unambiguous translation in Japanese. The Japanese version was developed by various translation and back-translation exercises and often a compromise had to be reached where a question needed to remain somewhat ambiguous in order to ensure it did not sound too formal. The lack of plural and singular words also led to ambiguity. One question asked how many generations of the participant's family had lived in Japan. In Japanese this question also meant 'how many generations of your family are still alive in Japan?' Because of the ambiguous nature of the language, questionnaires may not be a successful instrument. If one intends to include a Japanese sample, translation should be kept firmly in mind when developing the original questionnaire.

### Task observation

The observation sessions were particularly problematic for the Japanese. They felt uncomfortable speaking out loud about their thoughts and seemed to feel insecure because

they could not confer with others to reach a common opinion. As such, well-matched pairs or focus groups may be considered instead of the individual approach. The English also needed reassurance before feeling comfortable to talk out loud. Inhibition seemed stronger than for the Dutch and North-Americans. The Dutch tended to use sarcasm and humour during the sessions. This information often gets lost in translation or transcription and researchers may need to ensure answers are interpreted correctly.

### Interviews

The interview responses of the North Americans were often inconsistent with behaviour observed during the hands-on task. The researcher noticed participants were trying to give the 'right answer' rather than report their true findings in relation to the website. It may be that the North Americans felt slightly competitive when answering questions in a 'quiz-like' manner. As such it should be kept in mind that interview responses might not always gain reliable results.

Surprisingly, the Japanese participants seemed very comfortable in voicing positive as well as negative opinions of the website. It was expected that the individual setting would be intimidating and that they would feel uncomfortable about expressing opinions. However, the researcher noticed that when asked about their feelings in relation to the website they were quite happy to elaborate. Individually fulfilling a task with a particular goal had been something much less comfortable.

### CONCLUSIONS

The research described in this paper shows that one-on-one observation methods as commonly applied in usability testing may not be as appropriate for international users as it is for North-Americans. Other methods were also found to cause different complications for different cultures. More research is needed to investigate in detail the applicability of various user research methods. However, researchers should consider cultural differences when developing methods and instruments for cross-cultural user testing.

### REFERENCES

1. del Galdo, E., and Nielsen, J. (eds). *International User Interfaces*. Wiley, NY, 1996.
2. Fernandes, T. *Global Interface Design*. Academic Press, London, 1995.
3. Rubin, J. *Handbook of Usability Testing*. Wiley, NY, 1994.
4. Dumas, J. *Practical Guide to Usability Testing*. Intellect Books, 1999
5. Evers, V. *Cultural Aspects of User Interface Understanding*. Doctoral Thesis, the Open University, Milton Keynes, expected March 2002.