Usability Constructs: A Cross-Cultural Study of How Users and Developers Experience Their Use of Information Systems

Morten Hertzum[1], Torkil Clemmensen[2], Kasper Hornbæk[3], Jyoti Kumar[4], Qingxin Shi[2], and Pradeep Yammiyavar[4]

[1] Computer Science, Roskilde University, Roskilde, Denmark
mhz@ruc.dk
[2] Department of Informatics, CBS, Copenhagen, Denmark
{tc.inf,qs.inf}@cbs.dk
[3] Department of Computer Science, University of Copenhagen, Denmark
kash@diku.dk
[4] Indian Institute of Technology, Guwahati, India
{jyoti.k,pradeep}@iitg.ernet.in

Abstract. Whereas research on usability predominantly employs universal definitions of the aspects that comprise usability, people experience their use of information systems through personal constructs. Based on 48 repertory-grid interviews, this study investigates how such personal constructs are affected by two factors crucial to the international development and uptake of information systems: cultural background (Chinese, Danish, or Indian) and stakeholder group (developer or user). We find that for the user group frustrating and useful systems are experienced similarly, whereas for the developers frustrating systems are experienced similarly to easy-to-use systems. Looking at the most characteristic construct for each participant we find that Chinese participants use constructs related to security, task types, training, and system issues, whereas Danish and to some extent Indian participants make more use of constructs traditionally associated with usability (e.g., easy-to-use, intuitive, and liked). Further analysis of the data is ongoing.

Keywords: Cultural usability, Usage experiences, Repertory-grid technique.

1 Introduction

The concept of usability has been debated for decades [10, 16]. Most of this work, however, defines usability analytically or by reference to standards such as ISO 9241-11 [12]. Conversely, we know little about how people talk about their experiences with the systems they commonly use. Following Kelly [13] we take descriptions of such use experiences as indicative of the personal constructs people employ in relating to systems. By recognizing the personal nature of such usability constructs we seek to avoid unwarranted universalism and to explore how usability constructs are affected by two factors crucial to the international development and uptake of systems:

N. Aykin (Ed.): Usability and Internationalization, Part I, HCII 2007, LNCS 4559, pp. 317–326, 2007.

- *Cultural background*. The first aim of this study is to contribute to an elaboration of the cultural aspects of usability by investigating whether similarities and differences in people's usability constructs correlate with their cultural background. Cultural background is, in this study, taken to mean people's country of birth and residence. Though cultural usability is emerging as a topic [2, 9, 14], culture has typically not been considered at all in commonly accepted usability definitions.
- *Stakeholder groups*. The second aim of this study is to compare and contrast users' and developers' usability constructs. Any systematic differences in the usability constructs of these two stakeholder groups might impede user-developer communications about user requirements or system evaluations. Additionally, systematic differences may serve to elaborate and bridge between existing usability definitions.

To investigate the two factors empirically, we conduct repertory-grid interviews with users and developers with three different cultural backgrounds (Chinese, Danish, and Indian) and analyse the data descriptively and by means of principal-component analysis.

2 Related Work

While Barber and Badre [2] argue that users' cultural background can directly impact their performance using information technology (IT), the nature of this merging of culture and usability is presently far from clear. Research provides evidence that users' beliefs about their acceptance of systems and users' actual use of systems may be influenced by their cultural background. For example, Evers and Day [4] found that Chinese students attached more importance to perceived usefulness in forming an opinion about whether to accept a system, compared to Indonesian students who attached more importance to perceived ease of use. Honold [9] showed that washing machines were used quite differently in Indian and German households and that these differences led to fundamentally different user requirements. A prominent attempt at explaining the dimensions along which cultures differ is Hofstede's work [7], which identified five cultural dimensions: power distance, uncertainty avoidance, individualism/collectivism, masculinity/femininity, and long-term/short-term orientation. Hofstede's work has, for example, been introduced in HCI by Marcus and Gould [14] in relation to web-site design.

With respect to stakeholder groups, it is well-recognized that users and developers differ in manifold ways but, to our knowledge, no studies have systematically compared and contrasted how users and developers understand usability. Other stakeholder groups' understanding of usability have, however, been compared. Morris and Dillon [15] found that usability was not a central concern to managers responsible for making decisions about which IT systems to procure, but that it was a central concern for the end users. Moreover, managers and end users tended to conceptualize usability in different ways. To the managers, usability was predominantly a feature of the IT systems, such as 'having a point-and-click interface'. To the end users, usability was also dependent on the interactions among users, tasks, tools, and context. For example, one end user defined usability as "being able to use the software to perform the

tasks needed without excessive consultation" [15: p 253]. Holcomb and Tharp [8] had users rank the importance of the individual elements in a model of usability. Functionality was rated significantly more important than the six other elements of the model, namely consistency, user help, naturalness, user control, feedback, and minimal memorization. As the users had no option for extending the model with additional elements it was, however, not possible to say whether the model captured what the users considered to be the important elements of usability.

The repertory-grid technique, which is used in this study, originates from Kelly's personal-construct theory [13]. He rejected the idea that people perceive and make sense of their world by means of conceptions that exist independently of the individual person and instead proposed that people see their world through a set of personal constructs. These personal constructs are created over time in the course of people's interactions with their environment and express the dimensions along which a person can differentiate among objects and events. That is, each construct is defined by a similarity-difference dimension. Kelly [13] devised the repertory-grid technique to elicit personal constructs in the context of psychological counselling. The technique has subsequently been used successfully in interviews aiming to capture users' thoughts about IT products [1, 6, 17] and suggested for use in cross-cultural studies of information systems [11].

3 Method

To investigate the constructs people use to describe their experience of the information systems they use, we conducted repertory-grid interviews with people from two stakeholder groups (developers and users) and with three cultural backgrounds (Chinese, Danish, and Indian).

3.1 Participants

For each combination of stakeholder group and cultural background, we interviewed eight people. The Chinese participants lived and were interviewed in Beijing, the Danish participants in Copenhagen, and the Indian participants in Bangalore, Guwahati, Hyderabad, and Mumbai. Table 1 summarizes the 48 participants' gender, age, and IT experience. The participants had average to excellent English skills.

Table 1. Participant profiles

Group	*Gender*		*Age (years)*		*IT experience (years)*	
	Male	*Female*	*Mean*	*SD*	*Mean*	*SD*
Chinese developers	5	3	31.5	1.9	10.6	1.7
Chinese users	5	3	27.3	1.9	8.4	1.9
Danish developers	5	3	36.6	5.8	19.3	5.8
Danish users	5	3	36.8	6.2	16.9	3.6
Indian developers	8	0	29.6	1.7	9.9	2.5
Indian users	5	3	29.0	4.0	7.0	2.1

3.2 Procedure

Participants were interviewed individually by a person with a cultural background similar to their own. First, the study was described to the participant and the repertory-grid technique explained. Second, participants filled out a questionnaire about their background and signed an informed-consent form. Then, participants tried to elicit constructs with the repertory-grid technique on a couple of training tasks. After these preparatory steps, the actual repertory-grid interviews were conducted. They consisted of two steps: selection of systems and elicitation of constructs.

In selecting systems, the participant was asked to consider "*the array of computer applications you use for creating, obtaining, revising, managing, and communicating information and documents in the course of your day-to-day activities*." This included applications the participants use regularly but excluded applications they had only used once or twice and applications they merely know of. On this background participants were asked to select a system within each of six categories: my text processing system, my email, a useful system, an easy-to-use system, a fun system, and a frustrating system.

In eliciting constructs, the participant was successively presented with groups of three of the selected systems and asked: "*Can you think of some important way in which your personal experience using these three systems makes two of the systems alike and different from the third system?*" Having indicated the two similar systems, the participant wrote down a short phrase that told how these two systems were alike – the construct – and another short phrase that told how the third system differed – the contrast. Then, a seven-point rating scale was defined with this construct-contrast pair as its end points, and the participant rated all six systems according to this rating scale.

This procedure was repeated for all twenty combinations of three systems, in random order, or until the participant was unable to come up with a new construct for two successive combinations. The interviews were conducted in the participants' native language, if participants preferred that, or in English. Constructs and their contrasts were always recorded in English. In accordance with cultural customs, Danish and Indian participants received no compensation for their participation in the study while Chinese developers were paid 200RMB for their participation and Chinese users 50RMB. Each interview lasted about 1.5 hours.

3.3 Interviewer Preparations

The repertory-grid interviews were conducted by three of the authors. Three activities were performed to ensure that they conducted their interviews in the same way: First, we wrote an interview manual with step-by-step instructions about how to conduct the interviews. Second, each interviewer conducted a pilot interview. Third, we met before the pilot interviews to walk through a draft version of the interview manual and again after the pilot interviews to discuss experiences gained from the pilot interviews. The outcome of these preparations was the final version of the interview manual and a common understanding among the interviewers about how to conduct the interviews.

4 Results

We first present the participants' choice of systems and analyse the constructs used by individual participants. Next we analyse differences among systems, between stakeholder groups, and across participants' cultural backgrounds.

4.1 Participants' Choice of Systems

The 48 participants each selected six systems to be used in the elicitation of constructs. In the category 'my text processing system', 44 participants selected Microsoft Word; the remaining participants were divided on four additional systems. In the category 'my email', 20 participants selected Microsoft Outlook and seven participants selected Yahoo; the remaining participants were divided on seven additional systems. For the four other categories the participants selected a more mixed variety of systems. In the category 'a useful system' the most frequently selected system was Google (5 participants) and 36 additional systems were selected by one to four participants. In the category 'an easy-to-use system' Internet Explorer (5 participants) was the most frequent of a total of 30 different systems. In the category 'a fun system' three systems were selected by three participants (Google, Powerpoint, and Yahoo Messenger) and 32 additional systems by one or two participants. Finally, in the category 'a frustrating system' the most frequently selected system was Microsoft Excel (3 participants) and 42 additional systems were selected by one or two participants.

4.2 Constructs Used by Individual Participants

Participants reported an average of 13.8 constructs (SD = 3.6). The constructs varied much across individual participants in their level of abstraction, reference to personal experience, and relation to specific applications. Table 2 shows a summary of the most characteristic constructs as identified by principal-component analyses of individual grids. For each such analysis we selected the construct corresponding to the component that explained the largest amount of variance [5: pp 86-87, 3: p 14], for a total of 48 constructs.

Table 2. Participants' most characteristic construct. The table shows the most characteristic constructs that are shared by three or more of the 48 participants.

Most characteristic construct	*No. of participants*
Easy-to-use vs. Difficult	5
Work vs. Fun	5
Need for training vs. Walk-up-and-use	3
For myself vs. For the public	3
Simple vs. Complex	3

Across all 661 constructs, prominent kinds of construct relate to performance (e.g., 'Fast'), security (e.g., 'Easy to be affected by virus'), social issues (e.g., 'Communicate with other people'), frequency of use (e.g., 'Use everyday'), the context of use (e.g., 'Can use away from my desk'), the need to update and install programs (e.g., 'No need to update'), hedonic quality (e.g., 'Happy', 'Lot of fun to use'), aesthetics (e.g., 'Colourful interface'), and forgivingness (e.g., 'Insensitive to small mistakes').

4.3 Differences Among System Types

Fig. 1 shows the result of an individual differences multi-dimensional scaling on the six system types, across all 48 grids [5: p 99]. System types appear close together on the figure if participants rated them similarly on the rating scales defined by the construct-contrast pairs and far apart if participants rated them differently. The most noteworthy observation from this analysis is that the useful system and the frustrating system are close together, suggesting that participants rated these systems similarly. This observation is confirmed by an analysis of correlations of ratings among systems showing that ratings of frustrating systems are negatively correlated with ratings of all system types (r = -.14 to -.31, all ps < .001), except the useful system (r = .028, p > .4). This is not to say that frustrating systems are useful, but merely that usefulness does not indicate absence of frustration. For 25% of the 661 constructs, the ratings of the frustrating and the useful system are identical.

Fig. 1 also indicates that participants rate easy-to-use and fun systems similarly. Along one of the two dimensions in the multi-dimensional scaling easy-to-use and fun systems are also rated in opposition to useful and frustrating systems.

4.4 Differences Between Stakeholder Groups

Fig. 2 suggests that the two stakeholder groups conceptualize the systems differently. One difference is that for developers the frustrating system is close to the easy-to-use

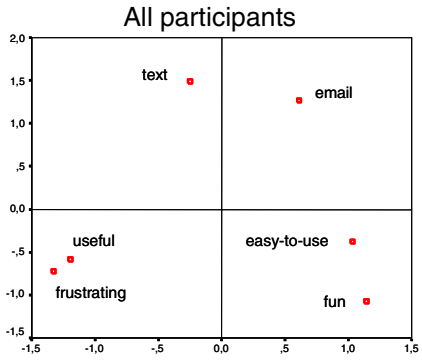

Fig. 1. Multi-dimensional scaling of system types based on data from all 48 participants. The stress value – an indicator of how well the scaling fits the raw data – for this scaling is .379.

system; this association is not found for the user group. Correlations of raw ratings show that easy-to-use and frustrating systems are not significantly correlated for developers (r = -.11, p > .05), but have a significant negative correlation for users (r = -.22, p < .001). For the user group we find a relation between the frustrating system and the useful system similar to that discussed in Section 4.3. An explanation of the difference between stakeholder groups for easy-to-use and frustrating systems may be that easy-to-use systems often cannot match the complexity of developers' work tasks and therefore resemble systems that cause developers frustrations. Another explanation may be that developers have higher or different standards for what constitutes an easy-to-use system. These explanations are merely tentative for three reasons: the dimensions of the plots in Fig. 2 are not easily comparable, the systems chosen as frustrating vary considerably across participants, and the constructs used by the two stakeholder groups may differ.

Another difference between the two stakeholder groups is that email seems to resemble text-processing systems for the developers, whereas for the user group email shares many of the properties of easy-to-use systems.

4.5 Differences Across Cultural Backgrounds

Fig. 3 shows a separate multi-dimensional scaling for participants with each of the three cultural backgrounds. From the diagrams it seems that systems for text and email are construed differently across cultures. In contrast to Danish and Indian participants, Chinese participants seem not to associate text-processing and email systems with each other (r = .008, p > .8), possibly reflecting a different role of email or issues associated with the support for writing Chinese. A further difference is that the fun system is associated with a different system for each of the participants' cultural backgrounds: for Chinese participants it is email, for Danish participants it is the easy-to-use system, and for Indian participants it is the useful system.

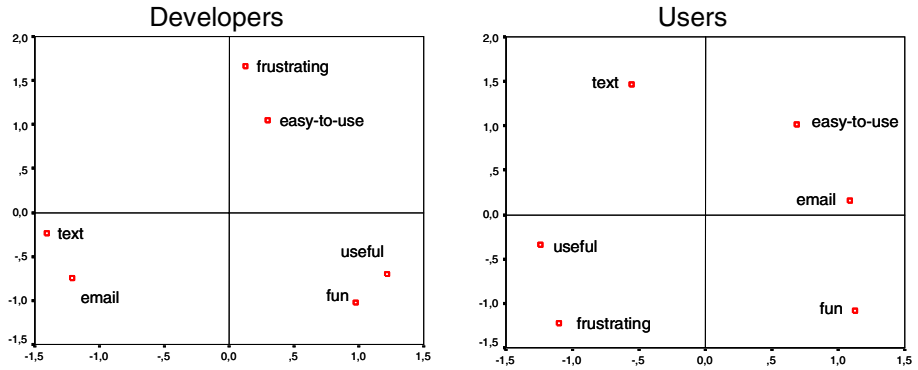

Fig. 2. Multi-dimensional scaling of system types: left panel is based on data from the 24 developers, and right panel is based on data from the 24 users. The stress values for these scalings are .385 and .311.

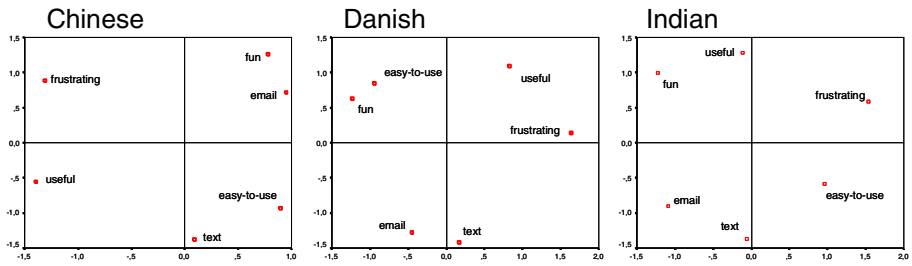

Fig. 3. Multi-dimensional scaling on participants' cultural background: left panel is based on data from the 16 Chinese participants, middle panel is based on data from the 16 Danish participants, and right panel is based on data from the 16 Indian participants. The stress values for these scalings are .319, .331, and .331.

Table 3 suggests that participants' cultural background influences which constructs they employ. Chinese participants have as their most characteristic construct a range of issues related to security, task types, training, and system issues. In contrast, Danish and to some extent Indian participants seem to mention more frequently aspects traditionally associated with usability (e.g., 'Easy-to-use', 'Intuitive', and 'Liked'). Eight (Danish) and six (Indian) of the most characteristic constructs are of this kind, as opposed to none of the constructs elicited by Chinese participants. Further, a distinction between work and leisure activities is more widely reported by Indian participants. Among all 661 constructs, however, the number of constructs that can be related unequivocally to this distinction are 15 (Indian), 11 (Chinese), and 12 (Danish).

Table 3. The most characteristic construct for each participant, divided onto cultural background. Some constructs have been slightly rephrased to be intelligible out of context.

Chinese participants	*Danish participants*	*Indian participants*
Often bring virus to computer	Experienced	Creative
Are used mostly by professionals	Stable and robust	Straight-forward
Daily use	Stand-alone program	Helps structuring
Used for email	Supports browsing	Natural way of use
Automatic installation	Give overview	Intuitively trustworthy
Use for programming	Context help	Complex product
Infrequent updating	Single supplier of application	Simpler
Use it when chatting	Simple	Stand alone application
Need internet connection	Easy-to-use	Just for relaxing
Can input information	Support numbers and figures	Help available
Can create something with applications	Easy-to-use	Entertainment
Need to use id	Intuitive	For work
Need training	Give focus	Recreation
Have many users	Process information	Liked
Need more memory	More complicated	Effective tools
Can use it first time	Creative	Related to public

5 Discussion and Conclusion

The participants in this study made use of a rich variety of constructs in talking about their experiences using IT systems. Following Kelly [13] these constructs, and their associated contrasts, define the dimensions along which participants perceive and are able to differentiate among usage experiences with different systems. Hence, the constructs can be seen as the building blocks of the participants' personal concepts of usability. In this sense the constructs stand in contrast to most definitions of usability, in which usability is defined analytically or with reference to standards like ISO 9241-11 [12]. An implicit assumption of these definitions is that they are valid across stakeholder groups and persons with different cultural backgrounds. Our analysis suggests that this assumption may not hold.

In this study, 48 participants made use of 661 construct-contrast pairs in describing how their experiences using some systems are alike and different from their experiences using other systems. Some of the constructs used by participants fit well with common definitions of usability, for example by emphasizing ease-of-use. Other constructs are well-known to human-computer interaction in that they describe use situations, the need for training, or frequency of use. However, a number of the elicited constructs are hard to reconcile with prevailing definitions of usability. For example, participants frequently mentioned issues of security – relating both to viruses and trustworthiness. The distinction between work and leisure is another example of a construct frequently employed by participants in distinguishing among systems but mostly glossed over in models of usability [e.g., 10, 12].

Some of the differences in the constructs employed by participants appear to be related to participants' cultural background and stakeholder group. A fun system is experienced similarly to email by Chinese participants, similarly to easy-to-use systems by Danish participants, and similarly to useful systems by Indian participants. The most characteristic construct for each participant provides further evidence for cultural differences in how the use of IT systems is experienced. Whereas traditional usability aspects, such as intuitiveness, are frequent among the most characteristic constructs of Danish and to some extent Indian participants, they are absent for the Chinese participants. This suggests cultural variation in the participants' concept of usability. In addition, developers seem to experience frustrating systems similarly to easy-to-use systems, whereas users experience frustrating systems similarly to useful systems. This adds to previous work by Morris and Dillon [15] and points toward possible sources of confusion in user-developer communication.

The present study has a number of limitations. First, the repertory-grid interviews were conducted by three interviewers. This may have introduced subtle differences in how interviews were conducted though we tried to avoid this through careful interviewer preparations. We chose against having the same interviewer for all interviews because it would mean that most or all participants would be interviewed by a person with a cultural background and native language different from their own. Second, some of the elicited constructs cannot readily be interpreted as aspects of the participants' experiences using the systems (e.g., 'Can have many windows'). However, in the absence of clear criteria for when to exclude a construct we included all constructs in the analysis. Third, part of our analysis is based on the most characteristic construct for each participant and, thus, disregards all additional constructs elicited by the participants.

Further analysis, including content analysis of the constructs, is ongoing.

Acknowledgements. This study was co-funded by the Danish Council for Independent Research through its support of the Cultural Usability project. We are grateful to the interviewees, who participated in the study in spite of their busy schedules.

References

1. Baber, C.: Repertory grid theory and its application to product evaluation. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry, pp. 157–165. Taylor & Francis, London (1996)
2. Barber, W., Badre, A.: Culturability: The merging of culture and usability. In: Proceedings of the Fourth Conference on Human Factors and the Web. AT&T Labs, Basking Ridge, NJ (1998). Available at: http://zing.ncsl.nist.gov/hfweb/att4/proceedings/barber/index.html
3. Bell, R.C.: Using SPSS to Analyse Repertory Grid Data. University of Melbourne, Melbourne, AU (1997). Available at: http://eu.wiley.com/legacy/wileychi/fransella/supp/gridspssman.doc
4. Evers, V., Day, D.: The role of culture in interface acceptance. In: Howard, S., Hammond, J., Lindgaard, G. (eds.) INTERACT '97: Proceedings of the IFIP TC13 International Conference on Human-Computer Interaction, pp. 260–267. Chapman and Hall, London (1997)
5. Fransella, F., Bell, R., Bannister, D.: A Manual for Repertory Grid Technique, 2nd edn. Wiley, New York (2004)
6. Hassenzahl, M., Wessler, R.: Capturing design space from a user perspective: The repertory grid technique revisited. International Journal of Human-Computer Interaction 12(3&4), 441–459 (2000)
7. Hofstede, G.: Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations. Sage, Thousand Oaks, CA (2001)
8. Holcomb, R., Tharp, A.L.: What users say about software usability. International Journal of Human-Computer Interaction 3(1), 49–78 (1991)
9. Honold, P.: Culture and context. An empirical study for the development of a framework for the elicitation of cultural influence in product usage. International Journal of Human-Computer Interaction 12(3&4), 327–345 (2000)
10. Hornbæk, K.: Current practice in measuring usability: Challenges to usability studies and research. International Journal of Human-Computer Studies 64(2), 79–102 (2006)
11. Hunter, M.G., Beck, J.E.: Using repertory grids to conduct cross-cultural information systems research. Information Systems Research 11(1), 93–101 (2000)
12. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. ISO, Genève, CH (1998)
13. Kelly, G.A.: The Psychology of Personal Constructs. Norton, New York (1955)
14. Marcus, A., Gould, E.W.: Crosscurrents: Cultural dimensions and global web user-interface design. ACM Interactions 7(4), 32–46 (2000)
15. Morris, M.G., Dillon, A.P.: The importance of usability in the establishment of organizational software standards for end user computing. International Journal of Human-Computer Studies 45(2), 243–258 (1996)
16. Shackel, B.: The concept of usability. In: Bennett, J., Case, D., Sandelin, J., Smith, M. (eds.) Visual Display Terminals: Usability Issues and Health Concerns, pp. 45–87. Prentice-Hall, Englewood Cliffs, NJ (1984)
17. Tan, F.B., Hunter, M.G.: The repertory grid technique: A method for the study of cognition in information systems. MIS Quarterly 26(1), 39–57 (2002)