*International Journal of Human-Computer Interaction*, vol. 27, no. 8, pp. 729-761 (2011)
Preprint version

# Personal Usability Constructs: How People Construe Usability across Nationalities and Stakeholder Groups

Morten Hertzum
Computer Science, Roskilde University, Roskilde, Denmark, mhz@ruc.dk

Torkil Clemmensen
Department of Informatics, CBS, Copenhagen, Denmark, tc.inf@cbs.dk

Kasper Hornbæk
Department of Computer Science, University of Copenhagen, Denmark, kash@diku.dk

Jyoti Kumar
Indian Institute of Technology, Guwahati, India, jyoti.k@iitg.ernet.in

Qingxin Shi
Department of Informatics, CBS, Copenhagen, Denmark, qs.inf@cbs.dk

Pradeep Yammiyavar
Indian Institute of Technology, Guwahati, India, pradeep@iitg.ernet.in

**Abstract**. Whereas the concept of usability is predominantly defined analytically, people relate to systems through personal usability 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 systems: nationality (Chinese, Danish, or Indian) and stakeholder group (developer or user). We find no significant overall difference across nationalities, but further analyses suggest that conventional usability aspects such as ease of use and simplicity are prominent for Chinese and Danish but not Indian participants and that a distinction between work and leisure-related communication is central to Chinese and Indian but not Danish participants. For stakeholder groups, we find a significant overall difference between developers and users. Unlike developers, users associate ease of use with leisure and, conversely, difficulty in use with work-relatedness. Further, users perceive usefulness as related to frustration and separate from ease of use, while developers construe usefulness, fun, and ease of use as related. In construing usability, participants make use of several constructs that are not part of prevailing usability definitions, including usefulness, fun, and security.

**Keywords:** Cultural usability, Stakeholder groups, Personal constructs, Usage experiences, Repertory-grid technique.

## 1 Introduction

The concept of usability is central to human-computer interaction and has been debated for decades (e.g., Bennett, 1984; Bevan, 1995, 2001; Eason, 1984; Hornbæk, 2006; Miller & Thomas, 1977; Shackel, 1984, 1991; Thomas & Macredie, 2002). Most of this work, however, defines usability analytically or by reference to standards such as ISO 9241 (1998). Comparatively less work has approached usability from the perspective of how people construe their experiences with the systems they commonly use. Following Kelly (1955), we expect people to employ a set of personal constructs in relating to systems and their use. Such personal usability constructs may enhance analytic definitions of usability by confirming core aspects of existing definitions, identifying recurrent concerns that are absent or under-recognized in existing definitions, and revealing unwarranted universalism that disregards variation in how different groups of people construe usability. We believe personal usability constructs can help anchor and contextualize usability definitions, which by themselves provide little in terms of arguments for their differences in what they include and exclude.

This study aims to explore the personal usability constructs people employ in talking about systems they use regularly. Such constructs are shaped by many factors, including people's behaviours, beliefs, values, professional backgrounds, contexts of use, and cultures. Some researchers seek to understand these factors through in-depth contextual investigations (e.g., Beyer & Holtzblatt, 1998; Button, 2000; Gaver, Dunne & Pacenti, 1999). Such investigations are invaluable to understand particular contexts and to drive the design of software. They help less, however, in generalizing about personal usability constructs and the factors that shape them. In contrast to contextual investigations, repertory-grid interviews (Kelly, 1955) are a technique for eliciting participants' personal constructs and may more easily be conducted so as to compare factors that shape personal usability constructs in a systematic manner. We therefore use repertory-grid interviews to investigate personal usability constructs empirically.

We are particularly interested in investigating how personal usability constructs are affected by two factors:

- *Nationality*. People of different nationality may construe usability differently depending on their use situations, preferences, and cultural backgrounds. We use the term nationality as a coarse, but straightforward way of differentiating groups of people, in comparison to terms like culture or ethnicity. The first aim of this study is to investigate whether similarities and differences in people's usability constructs owe to their nationality (viz., Chinese, Danish, and Indian). We find looking at nationality important for two reasons. First, though cultural usability is emerging as a topic (Day, 1998a, 1998b; del Galdo & Nielsen, 1996; Smith & Yetim, 2004), issues such as nationality and culture are not considered at all in commonly accepted usability definitions. Second, understanding if differences exist between people of different nationalities is crucial to the international development and uptake of systems.
- *Stakeholder group*. Any systematic differences in the usability constructs employed by different stakeholders in systems-development projects might impede communication and create confusion about, for example, user requirements and system evaluations. The second aim of this study is to compare and contrast users' and developers' usability constructs. In the empirical part of this study, users and developers are seen as general roles; that is, we will not be interviewing users of systems made by the interviewed developers. We find looking at users and developers important because they have markedly different stakes in systems-development projects, and differences in usability constructs seem particularly likely between these two stakeholder groups.

Differences in the usability constructs of different people, whether across nationalities or stakeholder groups, will have implications for researchers in terms of revealing biases and omissions in present usability definitions. The identification of such biases and omissions may, in turn, pinpoint areas that require the development of new or more sensitive methods for usability work. For practitioners, the identification of differences in the usability constructs of different people will point at concrete areas where misunderstandings are likely in user-developer communication and international systems development. Failure to appreciate such areas of misunderstanding about what constitutes a usable system may lead to uninformed prioritizations, misapplication of resources, flawed designs, and rejected systems.

## 2 Related work

Usability has been defined analytically in multiple and sometimes inconsistent ways. While these definitions prescribe the content of the concept of usability at a general level, a small number of studies describe elements specific to the usability concepts of people with different nationalities and from different stakeholder groups. We end this section by briefly introducing personal construct theory (Kelly, 1955).

### 2.1 Analytic usability definitions

Shackel (1984) defined usability as "the capability in human functional terms to be used easily (to a specified level of subjective assessment) and effectively (to a specified level of performance) by the specified range of users, given specified training and user support, to fulfill the specified range of tasks, within the specified range of environmental scenarios." He explicitly considered utility and likeability as related to, but distinct from, usability. Nielsen (1993) maintained the separation between usability and utility but included satisfaction (i.e., likeability) in his usability definition. ISO 9241 (1998) extended the concept further by also including part of utility in their definition of usability as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." Lastly, ISO/IEC 9126 (2001) reverted to a narrower concept of usability by defining it as "the capability of the software to be understood, learned, used and attractive to the user, when used under specified conditions" and explicitly stating that though functionality, reliability, and efficiency are related to usability they are excluded from the concept. These four influential usability definitions are analytic and vary substantially in the dimensions they include in and exclude from the concept of usability. The inclusions and exclusions do, however, not reflect differences in the groups of people targeted by the definitions. Rather, the included dimensions are general and the restriction in the scope of the definitions is made by a generalized reference to ”specified users”. While this restriction recognizes a possible impact of such factors as users’ nationality and stakeholder group, it says nothing about how and to what extent these factors might impact people’s usability constructs.

### 2.2 Effects of nationality on usability concepts

Research relevant to how users’ nationality affects their conception of usability mainly talks about cultural background. A likely reason for this is that nationality is an easy, though simplified, way to operationalize cultural background. Barber and Badre (1998) argue that users’ cultural background can directly impact their performance using information technology (IT). Presently, the nature of this merging of culture/nationality and usability is, however, far from clear.

Evers and Day (1997) found that Chinese students attached more importance to perceived usefulness in forming an opinion about whether to accept a system interface, compared to Indonesian students who attached more importance to perceived ease of use. Australian students seemed to be driven by neither perceived usefulness, nor perceived ease of use in forming their opinion about whether to accept the system interface. This suggests that across the three nationalities students held different perceptions of what made the interface acceptable. Leventhal, Teasley, Blumenthal, Instone, Stone, and Donskoy (1996) found that Non-Americans rated an interface higher in terms of sophistication and style than did Americans. The interface was designed specifically to appeal to a European audience, and three of the four rating scales relating to its sophistication and style concerned adjectives that had been particularly emphasized in its design. The study suggests that the adjectives (classical, restrained, suave, and the one not particularly emphasized in the design: bottle-like) and the appearance of the interface might reflect a European or Non-American outlook. Choi, Lee, and Kim (2006) found that Finnish, Japanese, and Korean users of mobile phone services differed in their preferences regarding various concrete elements of the interface, content, and information architecture of the services. Tractinsky (1997) found a higher correlation between beauty and ease of use among Israeli students who rated layouts of automatic teller machines (ATMs) than among the Japanese students in a previous instance of the study (Kurosu & Kashimura, 1995). The direction of the difference was contrary to Tractinsky’s hypothesis, leading him to the conclusion that current knowledge is insufficient as a basis for predicting how nationality influences usability issues.

Honold (1999, 2000) illustrated that differences in how people with different nationalities relate to technologies were not restricted to user perceptions but also affected how technologies were actually used. For example, Honold (1999) showed that following the purchase of a mobile phone the main learning strategy of Chinese users was to ask sales staff, friends, and acquaintances for assistance, whereas Chinese users who had had their phone for some time vehemently rejected asking questions as a strategy for solving problems relating to their use of their phone. For German users the written help material was the main source of information throughout the process of learning to use their mobile phones. This emphasizes that preferences specific to one nationality may, but need not, change as a process evolves. Marcus and Gould (2000) exemplify how web-site design differs across countries and discuss such differences in terms of Hofstede’s (2001) five cultural characteristics. In total, these studies provide evidence that users’ nationality may influence their beliefs about their acceptance of systems, their perception of system interfaces, and their actual use of systems.

### 2.3 Effects of stakeholder group on usability concepts

With respect to stakeholder groups, few studies have, to our knowledge, systematically compared and contrasted how different stakeholder groups construe usability. It is, however, well-recognized that users and developers differ in many ways (e.g., Jiang, Klein & Discenza, 2002; Landauer, 1995).

Boivie, Åborg, Persson, and Löfberg (2003) found that among developers usability was typically perceived to be the same as “a good user interface” and often as a system property added at a late stage in the development process. Some developers did, however, hold a broader view of usability, including the contents of the system and how it would affect users’ work. Contrasting developers and software-quality practitioners, Wilson and Hall (1998) found that developers perceived software usability in terms of a rather loose user-participation or customer-service approach, and that they experienced this approach to be in stark contrast to the standards-and-procedure-driven approach of software-quality practitioners.

Morris and Dillon (1996) 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 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”. None of the managers mentioned users, tasks, tools, or context as part of a definition of usability or as factors contributing to usability. To the users, usability was in part a matter of standard elements (e.g., “ease of use”) and system properties (e.g., "an intuitive interface with extensive on-line help"). However, many users' definitions of usability made explicit mention of interactions among users, tasks, tools, and context. For example, one user defined usability as "being able to use the software to perform the tasks needed without excessive consultation" (Morris & Dillon, 1996: p 253).

Holcomb and Tharp (1991) 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. Using a repertory-grid approach, Crudge and Johnson (2007) found that nearly all users' experience of information-retrieval systems also had affective elements relating to distraction, confusion, frustration, boredom, and overload.

### 2.4 Personal constructs and repertory grids

Personal construct theory (Kelly, 1955) rejects the idea that people perceive and make sense of their world by means of conceptions that exist independently of the individual person and instead proposes 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 differentiates among objects and events. Constructs are bipolar in that each construct comprises a similarity-difference dimension, which may for example define a construct as consisting of simple versus complex. This construct is different from the constructs simple versus powerful (in which simple is akin to powerless) and simple versus engaging (in which simple is akin to uninspiring). People differ substantially in the richness of their repertories of construct and, consequently, in the refinement of their ways of construing the world and informing their actions. A person's constructs are not necessarily consistent with each other, and they can be explicitly formulated or implicitly acted out.

To elicit people's constructs, Kelly (1955) devised the repertory-grid technique. While Kelly devised the technique for use in the context of psychological counselling it has subsequently been put to use in a range of contexts (Fransella, Bell & Bannister, 2004), including the evaluation of IT products (Baber, 1996; Hassenzahl & Wessler, 2000; Tan & Hunter, 2002). In such evaluations the repertory-grid technique provides a structured means of evaluating systems along dimensions deemed relevant by the involved users. This way of capturing users' thoughts about systems can also be seen as a method for studying the constructs people employ in talking about their use of systems. Indeed, Baber (1996) suggested the repertory-grid technique as a means of defining users' conceptions of usability. Furthermore, the repertory-grid technique has been suggested for use in cross-cultural studies of information systems (Hunter & Beck, 2000).

## 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 three nations (China, Denmark, and India) and two stakeholder groups (developers and users).

### 3.1 Participants

For each combination of nationality and stakeholder group, we interviewed eight people, for a total of 48 participants. All participants were citizens and residents in their country, and the participants as well as their parents had been raised in this country. On this basis, we consider the participants valid representatives of their nationality. The Chinese participants lived and were interviewed in Beijing, the Danish participants in Copenhagen, and the Indian participants in Bangalore, Guwahati, Hyderabad, or Mumbai. Table 1 summarizes the participants' gender, age, and IT experience. Developers had an average job experience as software developers of 6.1 ($SD$ = 1.5), 12.0 ($SD$ = 6.3), and 6.3 ($SD$ = 1.6) years for Chinese, Danish, and Indian developers, respectively. Thus, both developers and users had years of experience as representatives of their stakeholder group. The participants had average to excellent English skills, a qualification required for constructs to be recorded, but not necessarily elicited (see Section 3.2), in a uniform language.

To further characterize the participants, we asked them three questions about their use of information systems central to the repertory-grid interview and three general questions, adopted from Ceaparu, Lazar, Bessiere, Robinson, and Shneiderman (2004), about their association with information technology. Responses to all six questions consisted of ratings on seven-point rating scales. As recommended by Rosenthal and Rosnow (1991), we use analysis of variance on these ordinal data. Table 2 shows participants' responses. Participants differed significantly across nationalities in their use of text processing, $F(2, 46) = 7.04$, $p < .01$, and they approached a difference across stakeholder groups in their use of email, $F(1, 47) = 3.66$, $p = .06$. For the other questions, neither nationality nor stakeholder group yielded significant differences. Also, there were no significant interactions between nationality and stakeholder group for any of the six questions. On this basis, we consider the six groups of participant similar in most respects of relevance to this study, apart from the intentional difference in their nationality and stakeholder group.

### 3.2 Procedure

Participants were interviewed individually at their workplace, except one developer who was interviewed away from work. First, the study was described to the participant and the repertory-grid technique explained. Second, participants read and signed an informed-consent form and then filled out a questionnaire about their background. Third, participants were introduced to the repertory-grid technique and performed three to four training tasks to become familiar with the process of construct elicitation. After these preparations, the actual repertory-grid interviews were conducted. They consisted of three steps: selection of systems, elicitation of constructs, and rating of systems based on constructs.

Selection of systems had to be done by participants individually to ensure that they had experience using the systems. 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 used regularly but excluded applications they had only used once or twice and applications they merely knew 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. If a participant selected a system for a category but had already selected this system for another category, the participant was asked to select a different system. Thus, the selection process resulted in the selection of six different systems. Selection of elements (in our case, systems) from categories, as opposed to pre-selected elements, is common in repertory-grid studies (Fransella et al., 2004; Kelly, 1955; Tan & Hunter, 2002), and six elements are generally considered to provide sufficient variability for eliciting an exhaustive set of constructs (Baber, 2005; Hunter & Beck, 2000). We chose the six categories to balance inclusion of commonly used systems (e.g., text processing) against ensuring diversity in participants' experiences using the selected systems.

When the participant had selected a system in each of the six categories, the interview proceeded with the elicitation of constructs. 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 word or short phrase that told how these two systems were alike – the construct – and another word or short phrase that told how the third system differed – the contrast. If the construct and its contrast were unclear the interviewer would follow up by saying: "That is *one* way in which they are alike. Can you tell me how their being X [where X was the candidate construct] makes your personal experience of using these systems alike, and different from the third system?" The participant then clarified or changed the construct/contrast pair. This procedure of construct elicitation and follow-up questions was adopted from Kelly (1955).

When a construct/contrast pair had been successfully elicited, and before presenting the participant with the next group of three systems, the participant was asked to rate all six systems on a seven-point rating scale with the construct/contrast pair as its end points. This step ensures that all systems are reviewed relative to all construct/contrast pairs. While Kelly (1955) preferred a binary rating indicating that a construct either applied or did not apply, the use of rating scales with more than two points has subsequently become commonplace in repertory-grid studies (Fransella et al., 2004).

The steps of construct elicitation and system rating were 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. This was possible because all interviews were conducted by a person with the same nationality as the participant. Our reason for granting participants this choice of language was that the verbal skills of both participant and interviewer have been found important to successful repertory-grid interviews (Hassenzahl & Wessler, 2000). Constructs and their contrasts were always recorded in English. Any translation of constructs and contrasts into English was performed by the participant. In accordance with local customs, Danish and Indian participants received no compensation for their participation in the study while Chinese developers were paid RMB 200 for their participation and Chinese users RMB 50. Each interview lasted about 1.5 hours.

### 3.3 Interviewer preparations

The repertory-grid interviews were conducted by three of the authors. To ensure that they conducted their interviews in the same way, we first wrote an interview manual with step-by-step instructions about how to conduct the interviews and forms for recording systems, constructs, and ratings. The interview manual also included a participant profile to guide interviewers in recruiting participants. We met to walk through a draft version of the interview manual in preparation for a round of pilot interviews. Each interviewer conducted one pilot interview, and then we met again to discuss experiences gained from the pilot interviews and revise the interview manual accordingly. The outcome of these preparations was the final interview manual and a common understanding among the interviewers about how to conduct the interviews.

### 3.4 Data analysis

We analysed the repertory-grid data by means of qualitative content analysis and a variety of quantitative analyses. The two sets of analysis were independent in the sense that they were both made directly on all 661 elicited constructs; thus, noise in either the qualitative or quantitative analyses did not carry over to the other.

The content analysis involved two steps. First, two of the authors who had not been conducting repertory-grid interviews collectively made an affinity diagram of all 661 elicited constructs. This resulted in 51 categories, each described by a sentence. The categories were, in turn, organized into five kinds of construct. Second, to assess the reliability of the categories the three authors who had been conducting the interviews individually classified all constructs by assigning each construct to one category. We refer to these authors as the judges. Each judge was trained on a selection of 30% of the constructs, during which they twice received feedback on their classification. Different training sets were randomly selected for each judge. Some minimal changes were made to the descriptions of the categories as a result of the judges' classifications. Then each judge classified the 463 constructs not seen during training. Across the non-training constructs, the Kappa values for the level of agreement between the three judges and the categories were .64, .64, and .66, which according to Landis and Koch (1977) represents "substantial" agreement.

The quantitative analyses of the 661 constructs were based on generalized Procrustes analysis and principal component analysis. Generalized Procrustes analysis (Gower, 1975) is a multivariate analysis technique widely used to analyse for instance the experience of foods (Arnold & Williams, 1986), similarity of shapes (Goodall, 1991), perceptions of colour stimuli (Gains & Thomson, 1990), and dimensions of personality (Grice, Jackson & McDaniel, 2006). The basic idea of generalized Procrustes analysis is to help determine the consensus among a set of p points in k-dimensional space: in our case, among six systems and up to 20 construct/contrast pairs for each participant. The analysis uses translation, scaling, and rotation to arrive at a consensus configuration, a least-squares estimation of the agreement among the set of points. The consensus configurations for different groups may be compared, the relation between an individual's assessments and the consensus can be estimated, and the dimensions of the consensus configuration may be characterized. Our second analytic technique, principal component analysis, is used to project multidimensional data, such as those resulting from generalized Procrustes analysis, to fewer dimensions so that these new dimensions explain the largest amount of variance in the data. Each dimension can then be interpreted in terms of constructs that have high loadings on it. The use of principal component analysis to understand repertory grids is discussed in detail by Fransella et al. (2004) and Bell (1997). When reporting the results of principal component analyses we include only components with loadings equal to or above ± .7, as done by for instance Russell and Cox (2004). The analyses were performed in IdioGrid, a specialized software package for analysing repertory grids (Grice, 2002).

### 3.5 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 eight additional systems were selected by one to seven participants. 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.

Unsurprisingly, developers selected a larger proportion of technical systems (e.g., programming tools) than users. Conversely, multimedia systems (e.g., music and video applications) were selected more frequently by users. Danish participants selected slightly fewer systems intended for communication than did Chinese and Indian participants, and Chinese participants selected comparatively fewer Internet systems other than systems for communication. Apart from these minor differences, the six groups of participant selected a similar mix of different kinds of system.

# 4 Results

We first present the participants' choice of construct/contrast pairs. Next we analyse overall differences among participants, across nationalities, and between stakeholder groups. Finally, we analyse in more detail the constructs about use experience.

## 4.1 Constructs used by individual participants

Participants reported an average of 13.8 construct/contrast pairs (*SD* = 3.6), for a total of 661 pairs. One participant reported only 3 pairs and three participants only 8 pairs; five participants reported the maximum of 20 pairs.

The qualitative analysis of these constructs identified five kinds of construct: system characteristics (e.g., with or without a particular functionality), task or use-context characteristics (e.g., for work or leisure), use experience (e.g., time efficient), user characteristics (e.g., for specialist or general users), and other (4% of the constructs, mostly ones that were too unclear to classify). These kinds of construct each include a number of categories of similar constructs, for a total of 51 categories. Table 3 summarizes the kinds of construct.

For each kind of construct, Tables 4 to 7 list the categories of construct for that kind. The construct groups most frequently mentioned are those relating to work/leisure, communication, frequency of use, tailorability, text/graphics, internet connection, time efficiency, utility, and ease of use. Participants also frequently mention constructs that vary greatly depending on particularities of the system and its use. The categories containing these constructs are defined by, for instance, particular domains or applications (44 constructs), specific functionality (35 constructs), or specific tasks (22 constructs).

The literature on repertory grid analysis asserts that construct/contrast pairs that mostly receive extreme ratings (1 or 7 in our case) are of particular interest (Landfield & Cannell, 1988; Fransella et al., 2004, pp. 121-122). It has been argued that extreme ratings indicate more meaningful and superordinate constructs. Of the 661 construct/contrast pairs × 6 systems = 3966 ratings given by participants, there were 1990 (50%) extreme ratings. The constructs that most frequently received extreme ratings concern utility (8% of the extreme ratings), particular domains and applications (8%), communication (7%), work versus leisure (5%), internet connection (5%), other (5%), specialists versus non-specialist (4%), frequency of use (4%), privacy (4%), and support for learning (4%). These categories span all kinds of construct mentioned above, suggesting no clear trend in the kinds of construct that participants give extreme ratings.

A correlation of ratings among systems shows that ratings of frustrating systems are negatively correlated with ratings of all other system types ($r$ = -.14 to -.31, all $p$s < .001), except the useful system ($r$ = .028, $p$ > .4). These results also hold if we look at the Euclidian distances between elements, a measure that is insensitive to the orientation of construct/contrast pairs. The only other significant between-systems correlations after Bonferroni adjustments are between email and useful systems ($r$ = -.20) and between text-processing systems and fun systems ($r$ = -.17).

Four of the six systems that participants were asked to analyse were defined in terms of use experience (e.g., an easy-to-use system). Only 5% of the construct/construct pairs reused those terms and then almost always with a contrast generated by the participant and with a qualification (e.g., "Work related/Just fun"). The labelling of the system types thus seems to have had little direct influence on participants' construct/contrast pairs.

## 4.2 Overall differences among participants

The initial question for the data is whether either stakeholder group or nationality in a strong way helps differentiate among participants' ratings. A simple way to investigate this question is to analyse the Euclidian distance between participants' ratings. McEwan and Hallett (1990) described how to derive an assessor plot from the generalized Procrustes analysis. The idea is to subject the distances between participants' ratings to principal coordinate analysis (also known as classical multidimensional scaling). That analysis can be used to examine whether participants differ in any systematic way. Figure 1 shows such an assessor plot for the 48 participants.

As can be seen from the plot, no clear clustering results from neither participants' nationality, $\chi^2$ (3, *N* = 48) = 5.51, *p* > .4, nor their membership of a particular stakeholder group, $\chi^2$ (6, *N* = 48) = 2.73, *p* > .4. If we instead look at the difference between intra-nation and inter-nation distances for each participant, we still find no difference across nationalities (*t* = -.054, df = 47, *p* > .9). However, using this more sensitive test, we find a difference between stakeholder groups (*t* = -2.30, df = 47, *p* < .05). The average distance among participants within a stakeholder group is about 5% smaller than the average distance to participants in the other stakeholder group. Thus participants' ratings of their construct/contrast pairs display no differences between nationalities but a significant yet modest difference between stakeholder groups. We next turn to analysing the nature of the reported constructs.

### 4.3 Differences across nationalities

Table 3 shows the kinds of construct reported by participants from each of the three nations. Overall, constructs do not seem to be used with similar frequency among participants, $\chi^2$ (8, *N* = 661) = 19.99, *p* < .05. In particular, constructs about system characteristics are reported more frequently by Chinese participants (41% of their constructs) compared to Danish (33%) and Indian participants (24%). We find no one group of construct that Chinese participants particularly use and which could therefore explain the difference in frequency. Instead differences spread over groups of construct such as internet connection (9 constructs reported by Chinese participants, 3 by Danish participants, and 3 by Indian participants), local language (only Chinese), mentioning of specific functionality (Chinese: 18, Danish: 11, Indian: 6), and issues about system updates (Chinese: 9, Danish: 0, Indian: 1). Constructs about the task and use context are slightly more frequent among Indian participants (46%) compared to the other nationalities (Chinese: 33%, Danish: 38%). Indian participants seem to mention constructs relating to work or leisure more frequently (27 times) than Danish (17) and Chinese participants (14).

More subtle differences among participants may concern how constructs are used to distinguish between systems. We analyse differences between nationalities by principal component analysis of consensus configurations derived from generalized Procrustes analysis (see Figures 2 to 4). Next, we present each of these analyses in turn.

Figure 2 shows the two main dimensions of the Chinese participants' consensus configuration. These two dimensions account for 52% of the variance in the individual ratings. This amount (and those of the following analyses) is similar to several other studies (e.g., Elmore & Heymann, 1999; Russell & Cox, 2004). The first dimension on which Chinese participants construe systems concerns ease of use (or lack thereof, i.e., frustration). This interpretation is supported by the position of the easy-to-use and frustrating systems at each endpoint of the dimension. The dimension is related to construct/contrast pairs about ease of learning (e.g., "Easy to learn/Need more time to learn") and ease of use (e.g., "Easy to use/Difficult to use"); ease of use seems in particular to concern efficiency (e.g., "Longer time to finish ones work/Shorter"). This dimension is also related to the clarity of menu structures (e.g., "Menu structure is clear/Complex menu structure"). Only one of the 16 constructs that were classified as about use experience and loaded significantly on dimension one does not follow the interpretation of this dimension being primarily about ease of use. Among all 246 constructs reported by Chinese participants, ratings of the frustrating and the easy-to-use systems are negatively correlated (*r* = -.24, *p* < .001). Among the 49 constructs about use experience, 38 has an extreme rating of 1 or 7 for the frustrating and/or easy-to-use system and only 7 show the opposite direction from the interpretation suggested above.

The second dimension of Figure 2 seems related to communication, especially for private purposes, as opposed to work (e.g., "Contact with other people/Work tool"). This opposition seems to distinguish also the fun system from the useful system by placing them at each endpoint of the dimension: the useful system seems to be work related and the fun system to be communication related. The dimension furthermore reflects a distinction between general-purpose systems and software for work (e.g., "Used by public/Special tool for code") and whether systems are light-weight, requiring internet connection but with frequent updates, or heavy-weight, with a need to install software (e.g., "I must install before using them/Is provided by web site", "Long interval between new versions/At least one new version per year").

Note that Chinese participants distinguish the system they consider easy to use from that which is fun. These systems are distant on the consensus configuration in Figure 2, and the correlation between their ratings is not significant (*r* = .064, *p* > .3). In addition to the explanation given above relating to work, the difference between fun and easy-to-use systems may be illustrated by examining constructs that differ in their ratings for fun and easy-to-use systems. Here we see differences in ratings for efficiency (favouring the easy-to-use system) and with respect to aesthetics (favouring the fun system).

The consensus among Danish participants is summarized in Figure 3. The two dimensions in the figure account for 52% of the variance. As for Chinese participants, the first dimension of the Danish participants' consensus configuration is related to how easy a system is to use, for instance described with construct/contrast pairs on ease of use (e.g., "Easy-to-use/Always problems"), fun (e.g., "Happy to use/Annoying to use"), and user friendly (e.g., "User friendly/Non-user friendly"). Frequent use of systems (e.g., "Familiar/Do not use at the moment") and privacy (e.g., "Private use at work") seem also to play a role. Inflexibility (e.g., "Pre-fixed", "Non-controllable") appears the major opposite pole to these constructs.

The second dimension on which Danish participants differentiate systems is less clear. As dimension one, it includes simplicity and user friendliness. In contrast to dimension one, however, the second dimension includes likeability (e.g., "Likeable/Annoying") and stability (e.g., "Stable and robust/Unstable, breaks down"). The second dimension also seems related to whether a tool must be used or not (e.g., "Mandatory/Optional").

In contrast to the Chinese participants, easy-to-use and fun systems are close on the Danish participants' consensus configuration. Despite the proximity of these systems, they do not show a significant correlation in ratings ($r$ = -.011, $p$ > .8). So while they are similar in ratings along the two dimensions shown on Figure 3 they differ in ratings on the dimensions accounting for the remaining variation in the data. This may be due to constructs not directly about use experience, but for instance about differences in system types or tasks. If we consider just the constructs related to the use experience, we find a significant correlation between the easy-to-use and fun systems ($r$ = .37, $p$ < .01). These systems differ, however, with respect to most other kinds of construct. Another difference compared to the Chinese participants is that frustrating and fun systems are rated differently by the Danish participants ($r$ = -.352, $p$ < .001). Danish participants seem to consider frustration and fun as opposites, whereas Chinese participants do not ($r$ = -.098, $p$ >.1).

Figure 4 shows the results relating to Indian participants. The two dimensions in the figure account for 46% of the variance. The first dimension seems mostly related to work/leisure (e.g., "Work/Enjoyment"), specialist user (e.g., "For everybody/Very specialised, requires training"), and support for learning (e.g., "Flexible, easy to use, works quickly/requires time to learn"). The focus on the distinction between work and leisure accords with the already-mentioned observation that Indian participants report more construct/contrast pairs reflecting a work/leisure distinction than Danish and Chinese participants.

The second dimension of the Indian's consensus configuration also relates to the work-leisure distinction. In addition, it concerns construct/contrast pairs relating to communication (e.g., "Related to people/Related to files"), to aesthetics (e.g., "Black&white/Picturesque"), and to whether or not a system is engaging (e.g., "Entertaining, inspiration and creation/Same boring thing every time"). This dimension thus resembles the second dimension of the Chinese participants that also distinguished work from leisure activity. A similar importance of work-related constructs in separating systems is not found for the Danish participants.

Note that the consensus configuration for the Indian participants contains fewer references to conventional usability aspects. Only 10 of the 52 constructs about use experience are significantly correlated with the dimensions in Figure 4.

### 4.4 Differences between stakeholder groups

Table 3 shows that the five kinds of construct are used with similar frequency between stakeholder groups, $\chi^2$ (8, $N$ = 661) = 4.84, $p$ > .5. Contrary to nationality, membership of a stakeholder group does not influence the frequency with which participants use particular kinds of construct.

Figure 5 shows the consensus configuration for the interviewed developers. This configuration accounts for 48% of the variation in developers' use of constructs. The first dimension seems to concern activities involving other persons as opposed to work that is primarily individual. The positive dimension (work for others) is associated with high frequency of use (e.g., "Daily use/Infrequent use"). The frequency of use seems related to whether work is aimed at communication and sharing (e.g., "I can share my ideas with friends" or "Communication: official, family, friends, casual"). The negative end of this dimension (individual work) concerns whether the work is programming (e.g., "Programming/Office tool") or secondary work, including creative tasks and the use of task-specific applications. This dimension does not include conventional notions of usability, except for the mentioning of consistency (e.g., "Look and feel are same"), which is positively associated with text and email systems.

The second dimension of developers' consensus configuration concerns mainly frustration (or lack thereof). The frustrating system is at the positive endpoint of this dimension, whereas the negative endpoint is associated with constructs about likeability (e.g., "Happy to use/Annoying to use", "Likeable/Annoying"), efficiency (e.g., "Poor performance, slow over the net and hangs sometimes/Easy to use, fast"), and other desirable characteristics such as learnability and user friendliness. In contrast to the first dimension, the second dimension captures several aspects of conventional usability definitions.

We initially explored the data using multidimensional scaling (Hertzum, Clemmensen, Hornbæk, Kumar, Shi & Yammiyavar, 2007) and found easy-to-use and frustrating systems close for developers. Our findings here, using generalized Procrustes analysis, are different. Developers' consensus configuration centres easy-to-use systems (see Figure 5). The literature on interpreting graphical analysis of repertory grids suggests that this may be due to a lack of correlation with the dimensions of the map (e.g., Dijksterhuis & Punter, 1990). Here that would be surprising, given the many constructs with high loadings on the second dimension. Easy-to-use ratings, however, are only significantly correlated with email systems ($r = -.12$, $p < .05$); for the other five categories of system we find no significant correlations. Developers appear to rate easy-to-use systems high only on some constructs relating to the use experience, suggesting that ease of use comprises several independent dimensions to them. We return to the issue of dimensions of usability in the discussion.

Figure 6 shows the consensus configuration for users. It accounts for 50% of the variation in the constructs used. The first dimension concerns mainly work as opposed to leisure (e.g., "Work related/Just fun", "For playing movie/Do scientific analysis"). The systems related to leisure are used by the general public whereas the work systems are often specialized (e.g., "Could be used by public/Need special knowledge", "For everybody, enjoyment/Very specialized, requires training"). Coinciding with the distinction between leisure and work is a number of conventional usability aspects. These are oriented so that positive aspects correspond to leisure and negative to work, for instance as in "Easy-to-use/Difficult to use" and "I can use it when I first see it/I cannot use it sufficiently without training".

The second dimension of users' consensus configuration is fuzzy. It includes construct/contrast pairs about whether or not a system is important or mandatory (e.g., "Very important for my work/Not a necessity"), and on systems that help unwind (e.g., "Work related solutions/Solutions for emotional issues"). This dimension accounts for much less variation (18%) compared to the first dimension (32%).

For users we find an association between the frustrating system and the useful system. The consensus configuration places those systems in the same quadrant. The raw scores for users show that ratings of these two systems are identical for 27% of the constructs. A correlation between ratings for the frustrating and the useful system are positive ($r = .15$, $p < .01$), in contrast to the negative correlations between the frustrating system and the other four systems ($r$s from -.15 to -.32, all $p$s < .01). For developers, the correlation between ratings for the frustrating and useful systems is negative, but not significantly so ($r = -.11$, $p > .05$).

Users and developers also differ in how useful systems are related to easy-to-use and fun systems. Casual inspection of Figures 5 and 6 shows that useful, fun, and easy-to-use systems are close for developers, but more spread out for users. For users, correlations among ratings for the fun and useful systems are significantly negative ($r = -.15$, $p < .01$) as are correlations among useful and easy-to-use systems ($r = -.16$, $p < .01$). For developers these systems are not significantly correlated. Apparently, the notion of usefulness works differently in developers' and users' understanding of the systems they use.

### 4.5 Constructs characteristic of different systems

The constructs elicited allow us to investigate how participants construe use experience. The constructs most directly relating to conventional measures of usability are listed in Table 6, for a total of 143 constructs. In addition, four of the systems used to elicit constructs are defined in terms of use experience, namely systems that are easy-to-use, frustrating, fun, and useful. The construct/contrast pairs on which these systems are given extreme ratings of one or seven can help characterize those systems, as argued in Section 4.1. Below we analyse these constructs to give insight into use experience, independently of nationality and stakeholder group.

The easy-to-use systems are often associated with utility (10 extreme ratings, e.g., "Job-essential"), predictability (4 extreme ratings, e.g., "No unpleasant surprises/Unpredictable"), comfortableness (4 extreme ratings, e.g., "Can get pleasure from using it"), ease of installation (4 extreme ratings, e.g., "Small, easy to install/Large tool, install takes much time"), and non-frustration (3 extreme ratings, e.g., "Happy to use/Annoying to use"). Unsurprisingly, easy-to-use systems also receive many extreme ratings on construct/contrast pairs about ease of use (7 extreme ratings). It should be noted that the easy-to-use systems are not directly related to construct/contrast pairs about whether systems are simple or complex (4 ratings at one end of the scale, 3 at the other).

The frustrating systems are mainly characterized by a lack of comfort (4 extreme ratings), by not being easy-to-use (7 extreme ratings), by being non-inspiring (3 extreme ratings, e.g., "Interesting/Monotonous), and by being slow. Unsurprisingly, extreme ratings on construct/contrast pairs relating to frustration are frequent and consistent (8 extreme ratings). None of the construct/contrast pairs challenge the common understanding of frustration. As mentioned earlier, ratings for the frustrating system are also negatively correlated with ratings of the easy-to-use and fun systems.

For the fun system, the extreme ratings relate to ease of use (7 extreme ratings), whether a system is inspiring and interesting (6 extreme ratings), simple (6 extreme ratings), aesthetic (4 extreme ratings, e.g., "Beautiful interface/Earthly interface"), and comfortable (4 extreme ratings). As expected, the fun system is also related to construct/contrast pairs about fun/frustration (6 extreme ratings). Fun systems do not appear to be related to predictability as they are both rated as predictable and unpredictable (2 and 3 extreme ratings, respectively). Efficiency is also not related in any clear way to fun (5 constructs in either direction). Thus, construing a system as fun seems to be orthogonal to construing it in terms of efficiency.

The useful system is associated with ease of use (6 extreme ratings), fun (as opposed to frustration, 5 extreme ratings), and inspiration (4 extreme ratings). Note that usefulness is neither related in a clear way to efficiency, nor to simplicity. Furthermore, the notion of usefulness varies in the construct/contrast pairs mentioned: four pairs mention a lack of necessity to use, and four other pairs construe the useful system as essential and important.

Finally, let us mention some observations from the data on use experience that we find surprising in relation to the usability literature. First, several constructs discuss installing and updating software; a topic rarely discussed on its own in the usability literature (7 constructs, e.g., "Large tool, install takes much time/Small, easy to install"). Second, the infrequency of construct/contrast pairs relating to consistency—an often discussed notion in human-computer interaction—is also surprising. We find only three construct/contrast pairs about consistency (e.g., "Look and feel are same"). Third, users talk to a lesser degree than might be expected in terms of usability attributes concerning use experience. Only 22% of the construct/contrast pairs were classified as concerning use experience. This suggests that definitions of usability as consisting primarily of efficiency and satisfaction are inconsistent with participants' usability constructs.

# 5 Discussion

This study outlines a content-rich approach to studying usability and to characterizing empirically the dimensions along which people speak about usability. The participants in this study relate to systems they use through usability constructs that involve a variety of distinctions and thereby provide opportunities for enriching analytic usability definitions and for challenging studies suggesting that some stakeholder groups have rather crude perceptions of usability (e.g., Morris & Dillon, 1996).

## 5.1 Differences in usability constructs across nationalities

At the overall level the analysis of distances in the consensus configurations showed no difference between Chinese, Danish, and Indian participants. That is, the differences between nationalities were not larger than the variation within nationalities. This finding calls for caution in the interpretation of our more detailed results.

Conventional usability aspects such as ease of use, simplicity, and user friendliness are prominent for Danish participants in that these aspects are part of both dimensions in their consensus configuration. Similarly, the first dimension of their consensus configuration for Chinese participants concerns ease of use as opposed to frustration. Conventional usability aspects appear to play a minor role in how Indian participants construe systems. The easy-to-use system is, for example, close to the centre of the Indian participants' consensus configuration and is thus neither related to its first nor to its second dimension. This suggests that the prominence people attach to usability aspects such as ease of use may depend on their nationality. As a consequence, established usability definitions such as those of Nielsen (1993) and Shneiderman and Plaisant (2005) might display a regional or, possibly, cultural bias in the primacy they assign to ease of use.

Our study also suggests differences in the role communication and work play in participants' experiences of using systems. For Chinese and Indian participants a distinction between work and leisure constitutes the second dimension of their consensus configuration. In both cases leisure is partly construed in terms of communication, and the fun system is close to the leisure end of the dimension. For Danish participants neither work nor communication is part of the two dimensions of their consensus configuration. A further result of our study is that constructs referring to system characteristics are more common among Chinese participants whereas constructs referring to the task or use context are more common among Indian participants. This might suggest that Indian participants to a larger extent than Chinese and Danish participants construe systems in terms of the context in which the systems are set, whereas Chinese participants are more likely to perceive systems independently of their context. While this suggests an interesting difference in how Chinese and Indian participants primarily tend to perceive objects it is inconsistent with Nisbett's (2003) work on cultural cognition. According to Nisbett (2003), both Chinese and Indian participants should attend to the context in which systems are set, whereas Danish participants should attend more exclusively to the systems.

Like Nisbett, many other researchers use nationality as an operationalization of culture, sometimes (as in the present study) with the additional requirement that participants as well as their parents have been raised in the country. From a critical point of view, this operationalization tends to equate culture with national majority culture and to under-recognize within-nation variation in people's belief and value system. To avoid these problems it is necessary either to drop culture in favour of a more neutral designation such as nationality or to embrace culture through ethnographic and contextual approaches such as contextual inquiry (Beyer & Holtzblatt, 1998), cultural probes (Gaver et al., 1999), ethnography (Button, 2000), or another approach devised to collect rich empirical data about groups for the purpose of getting insights into their beliefs, values, and practices. From a more approving point of view, the requirement that participants as well as their parents have been raised in a country implies that participants have in their daily lives and through many years been exposed to belief and value systems represented in their country. On this basis participants can be assumed to share a number of beliefs and values with others of the same nationality, without necessarily sharing all their beliefs and values with everybody. This allows for some heterogeneity in the belief and value system that constitutes the cultural background of the people in a country.

In the present study we are cautious to extend our results from nationality to cultural background. On the contrary, the absence of an overall difference between nationalities may partly be caused by the existence of multiple cultural backgrounds within each country, resulting in variation within as well as across nationalities. Another explanation for the absence of an overall difference between nationalities may be that international applications such as Microsoft Word, used by 44 of the 48 participants, or other experiences common to the use of computers have resulted in an international "IT culture" that has a stronger influence on how participants construe usability than their nationality. If this explanation is correct, culture may have considerable influence on people's usability constructs but cannot be equated with nationality. This explanation is, however, weakened by the differences in our more detailed results for nationality.

### 5.2 Differences in developers' and users' usability constructs

The analysis of distances in the consensus configurations showed a significant difference between users and developers. That is, users were more similar to each other than to developers, and developers were more similar to each other than to users. The difference between developers and users was significant in spite of considerable variation within the two stakeholder groups. In other words, the differences in how users and developers construed usability were sufficiently strong to remain significant even though both stakeholder groups comprised Chinese, Danish, and Indian participants. On this basis, the differences between the two stakeholder groups seem important, and they add to previous work documenting differences in how users and managers perceive usability (Morris & Dillon, 1996).

The two main dimensions in the developers' consensus configuration concern activities involving others as opposed to primarily individual work and frustration as opposed to conventional usability aspects such as likeability and efficiency. Thus, conventional usability aspects appear to constitute a dimension of their own separate from a distinction between kinds of work. Conversely, the main dimension in the users' consensus configuration concerns work as opposed to leisure along with a distinction involving several conventional usability aspects. Thus, users construe conventional usability aspects along a dimension coinciding with their distinction between work and leisure and oriented so that work corresponds with the negative pole of the usability aspects. Contrary to developers, users appear to experience work-relatedness as involving systems that are difficult to learn and use, whereas ease of learning and use corresponds with leisure.

Developers and users also differ in how they construe usefulness. Users perceive useful and frustrating systems as similar in several respects. Overall users' ratings of useful and frustrating systems display a positive correlation, and for 27% of constructs users assign the same rating to useful and frustrating systems. Conversely, developers appear to construe frustrating systems differently from the other systems, including useful systems. Instead, developers rate useful systems similarly to fun and easy-to-use systems. The difference in how users and developers construe usefulness appears to be related to their different experience of work-related systems in that useful – and frustrating – systems are close to the work pole of the first dimension in the users' consensus configuration. This finding suggests that users consider usefulness a work-related construct and sees it as not only distinguishable but also different from conventional usability aspects such as easy to learn, easy to use, intuitive, and simple.

These results suggest differences in how stakeholders construe and talk about usability. These differences might point toward possible sources of confusion in user-developer communication. Developers are, for example, unlikely to perceive statements referring to a distinction between leisure and work as simultaneously conveying indications about whether a system is easy or difficult to learn and use. Users, on their part, perceive usefulness as closely related to frustration and clearly separate from ease of use; developers see usefulness, ease of use, and fun as close together. Thus, usefulness comes with very different connotations.

### 5.3 Usability constructs and extant usability definitions

In addition to the results relating to differences in usability constructs across participants' nationality and stakeholder group, we wish to relate the usability constructs to extant usability definitions. Rather than introducing a distinction between usability and some enclosing concept, such as user experience, we have adopted an inclusive notion of usability and discuss groups and kinds of construct within this inclusive notion. We chose against excluding certain constructs from usability because the criteria for doing it are vague and because we consider it more important to collect the constructs that matter to participants. The 51 groups of construct include constructs excluded from common analytic definitions of usability (e.g., ISO 9241, 1998; ISO/IEC 9126, 2001; Nielsen, 1993). Three of the authors made a post hoc categorization of the groups of construct by assigning them to either one of the three dimensions in the ISO 9241 (1998) definition of usability or to an ”other” category. As much as 22 groups of construct (43%) were unanimously assigned to the “other” category, indicating that they concern issues beyond effectiveness, efficiency, and satisfaction. We see this transcending of common usability definitions as a main contribution of the study, particularly because several of the novel constructs have high frequencies, which suggest that they are important.

We wish to note several issues in relation to common definitions of usability. First, conventional usability aspects such as ease of learning, ease of use, efficiency, and simplicity are important to how participants construe usability. Except for Indian participants, several of these aspects load on at least one of the dimensions of the consensus configurations.

Second, utility and usefulness seem important to participants. Evidence of this includes that construct/contrast pairs concerning utility often receive extreme ratings, that the useful system has a defining position near the work-related endpoint of the second dimension of the consensus configuration for Chinese participants, and that usefulness is perceived differently by developers and users. In addition, studies of technology acceptance (e.g., Venkatesh, Morris, Davis & Davis, 2003) find that perceived usefulness has a stronger and more lasting effect on people's perception of systems than perceived ease of use. Also, users in a study by Holcomb and Tharp (1991) ranked functionality, which is closely related to usefulness, the most important of seven aspects of usability. While usefulness is to some extent included in the ISO 9241 (1998) definition of usability, we are not aware of any attempt to tease apart dimensions of usefulness. This seems at odds with the many attempts to do so for usability.

Third, fun seems important to participants and is frequently contrasted with usefulness. For Chinese participants, Indian participants, and users, the fun system has a defining position near an endpoint of a dimension in the consensus configuration. These participants perceive the fun system as associated with leisure or unwinding and in opposition to work. The importance of fun accords with recent interest in fun (e.g., Blythe, Overbeeke, Monk & Wright, 2003) and joy of use (e.g., Hassenzahl, Beu & Burmester, 2001) and it goes beyond the more restricted satisfaction aspect of many usability definitions. The more restricted satisfaction aspect in, for example, the ISO 9241 (1998) definition of usability is more similar to participants' construct of likeability. While likeability coincide with usability aspects such as ease of learning and use, participants construe fun as a dimension different from these conventional usability aspects and instead associated with leisure.

Fourth, frustration seems to play a special role in how participants construe their use of systems. For all groups of participant the frustrating system has a defining position near one of the endpoints of a dimension in the consensus configuration. Interestingly, the opposite endpoint of these dimensions comprises both ease of use and leisure. While Chinese participants, Danish participants, and developers perceive frustration as opposed to ease of use, Indian participants perceive frustration as associated with work and in opposition to leisure, and users perceive frustration as opposed to both leisure and ease of use. Across nationalities and stakeholder groups frustration appears to capture a meaningful and important part of participants' experience with the systems they use. Conceptually frustration appears a clearly perceived aspect of usability, an aspect that on a practical level is something to be avoided. The high frequency and long accumulated duration of frustrating experiences during everyday system use (Ceaparu et al., 2004) provide further support for the importance of frustration in understanding and achieving usability.

Fifth, some elicited constructs are hard to reconcile with prevailing definitions of usability. For example, participants frequently mention 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 absent in definitions and models of usability. The usability of installation and updating – while probably not having a direct effect on models of usability – seems neglected, as many participants mention it as problematic.

These issues imply a need for extending usability work with methods more fully addressing usefulness and fun, which are currently secondary to efficiency and satisfaction (Hornbæk, 2006). Moreover, the constructs mentioned most frequently by participants include several not covered by current definitions, suggesting that some of the excluded constructs are central to how participants construe usability. For example, the most frequently mentioned group of construct is work versus leisure, which emphasizes a distinction that is being blurred in many other contexts. As the importance of this distinction varies across nationalities and stakeholder groups it appears to be an important area of attention in usability and usability work. The diversity of usability constructs raises a general issue of how to prioritize among them. This is a complex issue because correlations among constructs cannot be assumed (Frøkjær, Hertzum & Hornbæk, 2000) and may differ across, for example, nationalities. Yet, failure to prioritize the usability constructs important to users may ultimately result in system rejection.

### 5.4 Limitations

This study has four limitations that should be remembered in interpreting the results. First, the relationship between nationality and cultural background is unresolved. Though our selection of participants fulfils criteria adopted from research on cross-cultural differences, we have analysed our data in terms of the more neutral designation nationality. Extension of our findings to discussions of cultural differences should be done cautiously or not at all. Second, we interpret participants' answers to the question "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?" as their usability constructs. This results in an inclusive notion of usability. Others have defined usability more exclusively (e.g., ISO/IEC 9126, 2001; McGee, Rich & Dumas, 2004; Nielsen, 1993), and we acknowledge that several of the elicited constructs about the use experience may be considered beyond usability. Third, participants should have sufficient English skills to record their construct/contrast pairs in English. This requirement was harder to satisfy for Chinese participants than for Danish and Indian participants. Thus, Chinese participants may represent a smaller segment of their country, and it may have been harder for them to express their construct/contrast pairs in writing. When participants preferred it, constructs and contrasts were, however, elicited in participants' native language, alleviating the effect of the requirement to record construct/contrast pairs in English. Fourth, repertory grids are but one way to analyse personal usability constructs. One criticism that has been raised against repertory grids is that they rely on differentiation of elements (in our case, the six systems) and may, thereby, overlook construct/contrast pairs that are important to the participant but do not differentiate among the elements (Hassenzahl & Wessler, 2000). We aimed to counter this limitation by having participants select systems from six categories that ensured heterogeneity among the systems, but there is a need for complementing repertory grids with other methods in future work on personal usability constructs.

## 6 Conclusion

Usability is a central notion in human-computer interaction, but has mainly been defined analytically. We have used the repertory-grid technique to study how 48 participants made use of a rich variety of constructs in talking about their use of IT systems. Following personal construct theory, these constructs and their associated contrasts define the dimensions along which participants perceive and are able to differentiate among use experiences. This allows us to study usability on an empirical and personal basis.

Our study suggests few differences across nationalities. Rather, the yearlong use of similar IT applications may create a commonality that has a stronger effect on how participants construe usability than differences in their nationality. The role of systems for work and leisure, however, appears to depend on participants' nationality, as does the importance of conventional usability aspects, such as ease of use and simplicity. Differences between users and software developers seem more pronounced. Contrary to developers, users relate ease of use with leisure and frustration with work. Developers and users also differ in their use of notions such as usefulness. Our study raises several questions about existing models of usability because they do not capture constructs or relationships that our data suggest important: what does usefulness mean, what is the difference between usability in work and leisure, what is the relation between fun and usability, and what does the seeming importance of frustration to participants imply for models of usability? On a more general level, our results indicate that usability constructs vary a lot and, thereby, call for corroboration and further investigation into sources of the variation. For practitioners, our study suggests that stakeholder differences in usability constructs may affect user-developer communication; differences in nationality may impact the relative importance of usability aspects. For usability research, the findings illustrate one way of investigating how people construe usability, and they challenge a number of conventional assumptions about what constitutes usability. For both researchers and practitioners, it appears an important challenge to find better indicators of cultural background than nationality.

## Acknowledgements

This study was co-funded by the Danish Council for Independent Research (DCIR) through its support of the Cultural Usability project. We thank James W. Grice, Oklahoma State University, for help on IdioGrid and on analysis of repertory grids more generally. We also thank the interviewees, who participated in the study in spite of their busy schedules.

## References

Arnold, G.M., & Williams, A.A. (1986). The use of generalised Procrustes techniques in sensory analysis. In J.R. Piggott (Ed.), *Statistical Procedures in Food Research* (pp. 223-254). London: Elsevier Applied Science.

Baber, C. (1996). Repertory grid theory and its application to product evaluation. In P.W. Jordan, B. Thomas, B.A. Weerdmeester, & I.L. McClelland (Eds.), *Usability Evaluation in Industry* (pp. 157-165). London: Taylor & Francis.

Baber, C. (2005). Repertory grid for product evaluation. In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), *Handbook of Human Factors and Ergonomics Methods* (pp. 31.1-31.7). Boca Raton, FL: CRC Press.

Barber, W., & Badre, A. (1998). Culturability: The merging of culture and usability. In *Proceedings of the Fourth Conference on Human Factors and the Web*. Basking Ridge, NJ: AT&T Labs. Available at: http://zing.ncsl.nist.gov/hfweb/att4/proceedings/barber/index.html.

Bell, R.C. (1997). *Using SPSS to Analyse Repertory Grid Data*. Melbourne, AU: University of Melbourne. Available at: http://eu.wiley.com/legacy/wileychi/fransella/supp/gridspssman.doc.

Bennett, J.L. (1984). Managing to meet usability requirements: Establishing and meeting software development goals. In J. Bennet, D. Case, J. Sandelin, & M. Smith (Eds.), *Visual Display Terminals: Usability Issues and Health Concerns* (pp. 161-184). Englewood Cliffs, NJ: Prentice-Hall.

Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal*, 4(2), 115-130.

Bevan, N. (2001). International standards for HCI and usability. *International Journal of Human-Computer Studies*, 55(4), 533-552.

Beyer, H., & Holtzblatt, K. (1998). *Contextual Design: Defining Customer-Centered Systems*. San Francisco, CA: Morgan Kaufmann.

Blythe, M.A., Overbeeke, K., Monk, A.F., & Wright, P.C. (Eds.). (2003). *Funology: From Usability to Enjoyment*. Dordrecht, NL: Kluwer.

Boivie, I., Åborg, C., Persson, J., & Löfberg, M. (2003). Why usability gets lost or usability in in-house software development. *Interacting with Computers*, 12(4), 623-639.

Button, G. (2000). The ethnographic tradition and design. *Design Studies*, 21(4), 319-332.

Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J., & Shneiderman, B. (2004). Determining causes and severity of end-user frustration. *International Journal of Human-Computer Interaction*, 17(3), 333-356.

Choi, B., Lee, I., & Kim, J. (2006). Culturability in mobile data services: A qualitative study of the relationship between cultural characteristics and user-experience attributes. *International Journal of Human-Computer Interaction*, 20(3), 171-206.

Crudge, S.E., & Johnson, F.C. (2007). Using the repertory grid and laddering technique to determine the user's evaluative model of search engines. *Journal of Documentation*, 63(2), 259-280.

Day, D. (Ed.). (1998a). Shared values and shared interfaces [special issue]. *Interacting with Computers*, 9(3).

Day, D. (Ed.). (1998b). Shared values and shared interfaces 2 [special issue]. *Interacting with Computers*, 9(4).

del Galdo, E.M., & Nielsen, J. (Eds.). (1996). *International User Interfaces*. New York: Wiley.

Dijksterhuis, G., & Punter, P. (1990). Interpreting generalized Procrustes analysis 'analysis of variance' tables. *Food Quality and Preference*, 2(4), 255-265.

Eason, K.D. (1984). Towards the experimental study of usability. *Behaviour & Information Technology*, 3(2), 133-143.

Elmore, J.R., & Heymann, H. (1999). Perceptual maps of photographs of carbonated beverages created by traditional and free-choice profiling. *Food Quality and Preference*, 10(3), 219-227.

Evers, V., & Day, D. (1997). The role of culture in interface acceptance. In S. Howard, J. Hammond, & G. Lindgaard (Eds.), *INTERACT '97: Proceedings of the IFIP TC13 International Conference on Human-Computer Interaction* (pp. 260-267). London: Chapman and Hall.

Fransella, F., Bell, R., & Bannister, D. (2004). *A Manual for Repertory Grid Technique. Second Edition*. New York: Wiley.

Frøkjær, E., Hertzum, M., & Hornbæk, K. (2000). Measuring usability: Are effectiveness, efficiency, and satisfaction really correlated? In *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems* (pp. 345-352). New York: ACM Press.

Gains, N., & Thomson, D. (1990). The relation of repertory grid/generalized Procrustes analysis solutions to the dimensions of perception: Application to Munsell color stimuli. *Journal of Sensory Studies*, 5(3), 177-192.

Gaver, B., Dunne, T., & Pacenti, E. (1999). Cultural probes. *ACM Interactions*, 6(1), 21-29.

Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2), 285-339.

Gower, J. (1975). Generalized Procrustes analysis. *Psychometrika*, 40(1), 33-51.

Grice, J. (2002). Idiogrid: Software for the management and analysis of repertory grids. *Behavior Research Methods, Instruments, & Computers,* 34(3), 338-341.

Grice, J., Jackson, B., & McDaniel, B. (2006). Bridging the idiographic-nomothetic divide: A follow-up study. *Journal of Personality*, 74(4), 1191-1218.

Hassenzahl, M., Beu, A., & Burmester, M. (2001). Engineering joy. *IEEE Software*, 18(1), 70-76.

Hassenzahl, M., & Wessler, R. (2000). Capturing design space from a user perspective: The repertory grid technique revisited. *International Journal of Human-Computer Interaction*, 12 (3&4), 441-459.

Hertzum, M., Clemmensen, T., Hornbæk, K., Kumar, J., Shi, Q., & Yammiyavar, P. (2007). Usability constructs: A cross-cultural study of how users and developers experience their use of information systems. In N. Aykin (Ed.), *Proceedings of the HCI International Conference*. Lecture Notes in Computer Science No. 4559 (pp. 317-326). Berlin: Springer-Verlag.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations. Second Edition*. Thousand Oaks, CA: Sage.

Holcomb, R., & Tharp, A.L. (1991). What users say about software usability. *International Journal of Human-Computer Interaction*, 3(1), 49-78.

Honold, P. (1999). Learning how to use a cellular phone: Comparison between German and Chinese users. *Technical Communication*, 46(2), 196-205.

Honold, P. (2000). 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.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, 64(2), 79-102.

Hunter, M.G., & Beck, J.E. (2000). Using repertory grids to conduct cross-cultural information systems research. *Information Systems Research*, 11(1), 93-101.

ISO 9241 (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability*. Genève, CH: International Organization for Standardisation.

ISO/IEC 9126 (2001). *Software engineering – product quality – Part 1: Quality model*. Genève, CH: International Organization for Standardisation.

Jiang, J.J., Klein, G., & Discenza, R. (2002). Perception differences of software success: Provider and user views of system metrics. *Journal of Systems and Software*, 63(1), 17-27.

Kelly, G.A. (1955). *The Psychology of Personal Constructs*. New York: Norton.

Kurosu, M., & Kashimura, K. (1995). Determinants of apparent usability. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (pp. 1509-1513). Los Alamitos, CA: IEEE Press.

Landauer, T.K. (1995). *The Trouble with Computers: Usefulness, Usability, and Productivity*. Cambridge, MA: MIT Press.

Landfield, A.W., & Cannell, J.E. (1988). Ways of assessing functionally independent construction, meaningfulness, and construction in hierarchy. In J.C. Mancuso & M.L. Shaw (Eds.), *Cognition and Personal Structure* (pp. 67-90). New York: Praeger.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.

Leventhal, L., Teasley, B., Blumenthal, B., Instone, K., Stone, D., & Donskoy, M.V. (1996). Assessing user interfaces for diverse user groups: Evaluation strategies and defining characteristics. *Behaviour & Information Technology*, 15(3), 127-137.

Marcus, A., & Gould, E.W. (2000). Crosscurrents: Cultural dimensions and global web user-interface design. *ACM Interactions*, 7(4), 32-46.

McEvan, J., & Hallet, E. (1990). *A guide to the use and interpretation of generalized procrustes analysis*. Technical Manual No 30. Gloucestershire, UK: Campden & Chorleywood Food Research Association.

McGee, M., Rich, A., and Dumas, J. (2004). Understanding the usability construct: User-perceived usability. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 907-911). Santa Monica, CA: HFES.

Miller, L.A., & Thomas, J.C. (1977). Behavioral issues in the use of interactive systems. *International Journal of Man-Machine Studies*, 9(5), 509-536.

Morris, M.G., & Dillon, A.P. (1996). 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.

Nielsen, J. (1993). *Usability Engineering*. Boston, MA: Academic Press.

Nisbett, R.E. (2003). *The Geography of Thought – How Asians and Westerners Think Differently – and Why*. London: Brealey.

Rosenthal, R., and Rosnow, R.L. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. Second Edition. Boston, MA: McGraw-Hill.

Russel, C.G., & Cox, D.N. (2004). Understanding middle-aged consumers' perceptions of meat using repertory grid methodology. *Food Quality and Preference*, 15(4), 317-329.

Shackel, B. (1984). The concept of usability. In J. Bennett, D. Case, J. Sandelin, & M. Smith (Eds.), *Visual Display Terminals: Usability Issues and Health Concerns*. (pp. 45-87). Englewood Cliffs, NJ: Prentice-Hall.

Shackel, B. (1991). Usability – context, framework, definition, design and evaluation. In B. Shackel & S. Richardson (Eds.), *Human Factors for Informatics Usability* (pp. 21-37). Cambridge, UK: Cambridge University Press.

Smith, A., & Yetim, F. (Eds.). (2004). Global human-computer systems: Cultural determinants of usability [special issue]. *Interacting with Computers*, 16(1).

Shneiderman, B., & Plaisant, C. (2005). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Fourth Edition. Boston, MA: Addison-Wesley.

Tan, F.B., & Hunter, M.G. (2002). The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*, 26(1), 39-57.

Thomas, P., & Macredie, R.D. (Eds.). (2002). The new usability [special issue]. *ACM Transactions on Computer-Human Interaction*, 9(2).

Tractinsky, N. (1997). Aesthetics and apparent usability: Empirically assessing cultural and methodological issues. In *Proceedings of the CHI'97 Conference on Human Factors in Computing Systems* (pp. 219-330). New York: ACM Press.

Venkatesh, V., Morris, M.G., Davis, G.B., & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.

Wilson, D.N., & Hall, T. (1998). Perceptions of software quality: A pilot study. *Software Quality Journal*, 7(1), 67-75.

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 |

Table 2. Participants' ratings of their use of and association with information technology, *N* = 48 participants.

| *Group* | *Q1* | | *Q2* | | *Q3* | | *Q4* | | *Q5* | | *Q6* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* |
| Chinese developers | 6.9 | .4 | 7.0 | .0 | 7.0 | .0 | 6.3 | .5 | 5.5 | 1.6 | 3.6 | 2.0 |
| Chinese users | 6.8 | .5 | 6.8 | .7 | 6.8 | .7 | 6.1 | .6 | 5.5 | 1.3 | 3.1 | 1.4 |
| Danish developers | 5.5 | 1.7 | 6.9 | .4 | 7.0 | .0 | 5.9 | 1.0 | 6.4 | .5 | 5.4 | 1.3 |
| Danish users | 6.1 | 1.5 | 6.8 | .5 | 6.8 | .5 | 5.9 | 1.5 | 6.0 | .9 | 3.9 | 2.2 |
| Indian developers | 5.3 | 1.4 | 6.8 | .5 | 6.9 | .4 | 6.0 | .9 | 6.3 | .7 | 3.4 | 1.3 |
| Indian users | 5.0 | 1.6 | 6.3 | 1.4 | 6.3 | 1.4 | 5.5 | 1.2 | 5.5 | .9 | 4.3 | 1.7 |

Notes:
Q1: I use text processing (1: never – 7: every day)
Q2: I use the web (1: never – 7: every day)
Q3: I use email (1: never – 7: every day)
Q4: How sufficient is your computer hardware and software for the work you need to do (1: not at all – 7: very)
Q5: Overall, computers make we feel (1: very uncomfortable – 7: very comfortable)
Q6: When you run into a problem on the computer or an application you are using, do you feel (1: anxious – 7: relaxed/indifferent)

Table 3. Overview of kinds of construct.

| *Kind* | *All* | *Nationality* | | | *Stakeholder group* | |
|---|---|---|---|---|---|---|
| | | Chinese | Danish | Indian | User | Developer |
| System characteristics | 218 | 100 | 65 | 53 | 110 | 108 |
| Task or use context | 256 | 81 | 76 | 99 | 144 | 112 |
| Use experience | 143 | 49 | 42 | 52 | 71 | 72 |
| Kind of user | 15 | 7 | 2 | 6 | 7 | 8 |
| Other | 29 | 9 | 13 | 7 | 11 | 18 |
| All | 661 | 246 | 198 | 217 | 343 | 318 |

Table 4. Constructs related to system characteristics.

| *Construct group* | *N* | *Explanation* |
|---|---|---|
| Specific functionality | 35 | Concerns systems with a specific functionality or requirement (e.g., search, history, memory) |
| Tailorability | 23 | Concerns the flexibility of systems, in particular the degree to which they may be tailored |
| Generic vs specific tasks | 22 | Distinguishes systems that work across domains/tasks from systems that are specific to a certain domain/task |
| Text vs graphics | 18 | Distinguishes systems using mainly text from those using graphics or visual contents |
| Internet connection | 15 | Systems that require internet or network connection to work correctly |
| Need for installation | 12 | Distinguishes pre-installed systems (or those that require no installation) from those that must be installed to work |
| Installation vs stand alone | 11 | Distinguishes systems that work on their own from systems that are integrated or a subsystem |
| System updates | 10 | Concerns the frequency and nature of system updates such as new versions |
| License | 8 | Systems that are free versus those that require a license |
| Reliability | 7 | Distinguishes systems that are reliable (stable) from those which are not |
| User initiated vs automatic | 7 | Distinguishes systems or functions that are automatically available or always enabled from applications/functions the user must initiate |
| Multiple vs single | 7 | Distinguishes systems that can work on multiple documents or windows from systems that cannot |
| Microsoft | 7 | Systems that are made by Microsoft as opposed to other vendors |
| Keyboard vs mouse | 6 | Distinguishes systems where a keyboard is used from those operated mainly by mouse |
| Virus | 6 | Systems that easily get infected by virus as opposed to those that do not |
| Popularity | 5 | Concerns the uptake or popularity of systems |
| Overview | 5 | Concerns whether or not a system provides an overview |
| Local language | 4 | Concerns whether or not systems provide support for a local language, e.g. Chinese |
| Numbers vs words | 4 | Systems that use or concern numbers as opposed to using mainly words |
| Alternative suppliers | 4 | Systems that are available from alternative suppliers or in alternative versions |
| Up-to-date | 2 | Concerns whether or not a system's content is up-to-date |

Note: *N* indicates the number of constructs in each construct group.

Table 5. Constructs related to the task or use context.

| *Construct group* | *N* | *Explanation* |
|---|---|---|
| Work vs leisure | 58 | Systems related to work/business as opposed to leisure/private concerns |
| Particular domains or applications | 44 | Concerns systems that provide support for programming, planning, courses, presentation, literature, and other specific domains and applications |
| Communication | 41 | Systems that support communication with others; often in contrast to systems for manipulating contents |
| Frequency of use | 27 | Systems that are used frequently or with which the user is familiar as opposed to those that are rarely used |
| Support for learning | 24 | Systems that support learning, in particular through training, exploration, and trial-and-error |
| Use vs produce | 14 | Distinguishes whether a system supports using information/services or producing content |
| Creativity | 14 | Systems that support being creative and getting ideas |
| Modifiability | 10 | Distinguishes systems where users may modify the content (save/write/store) from those that cannot be modified |
| Privacy | 8 | Issues relating to passwords, pin codes, and other privacy matters |
| Unwind | 7 | Systems that actively help users unwind and relax as opposed to systems that do not or are stressful |
| Push vs pull of information | 5 | Distinguishes systems that require the user to request/pull information from systems that push information |
| Availability of system | 4 | Systems that are available only at one particular place as opposed to everywhere |

Note: *N* indicates the number of constructs in each construct group.

Table 6. Constructs related to use experience.

| *Construct group* | *N* | *Explanation* |
|---|---|---|
| Time efficiency | 19 | Distinguishes slow from fast systems |
| Utility | 19 | Distinguishes systems that are seen as useful, urgent to use, or of utility to the user from those which are not |
| Ease of use | 17 | Distinguishes easy-to-use and helpful systems from difficult-to-use systems |
| Simplicity | 14 | Distinguishes systems that are simple (e.g., have few features) from those which are complex |
| Fun vs frustrating | 11 | Distinguishes systems that are fun or likeable from those which are frustrating or annoying |
| Aesthetics | 10 | Concerns the visual appeal of a system's interface |
| Comfortable | 9 | Concerns whether a system is pleasant to use, i.e. comfortable, friendly, and affective |
| Inspiring and interesting | 7 | Concerns systems that are inspiring, engaging, or interesting to the user |
| Predictable | 7 | Systems that are predictable as opposed to systems that surprise |
| Convenience of installation | 7 | Distinguishes systems that are easy to install and require few system resources from systems that are hard to install and require many system resources |
| Clear menu structure | 6 | Concerns whether or not a system's menu structure is easy to understand |
| Intuitive vs focused effort needed | 5 | Concerns systems that are intuitive to use as opposed to systems that require a focused effort |
| Consistency | 3 | Concerns whether or not systems are consistent |
| Powerful | 3 | Concerns powerful and expressive systems |
| Efficiency | 3 | Distinguishes systems or operations that require few steps from those which require many steps |
| Trust | 3 | Concerns how well a user trusts a system or how transparent its interface is |

Note: *N* indicates the number of constructs in each construct group.

Table 7. Constructs related to the kind of user

| *Construct group* | *N* | *Examples* |
|---|---|---|
| Specialists vs non-specialists | 15 | Distinguishes systems that require specialist/knowledgeable users (primarily technically knowledgeable) from systems that make no such requirements |

Note: *N* indicates the number of constructs in the construct group.

Figure 1. Assessor plot of the 48 participants. Participants' nationality is indicated by colour (Danish: light grey, Chinese: dark grey, Indian: black), their stakeholder group by shape (developer: square, user: circle).

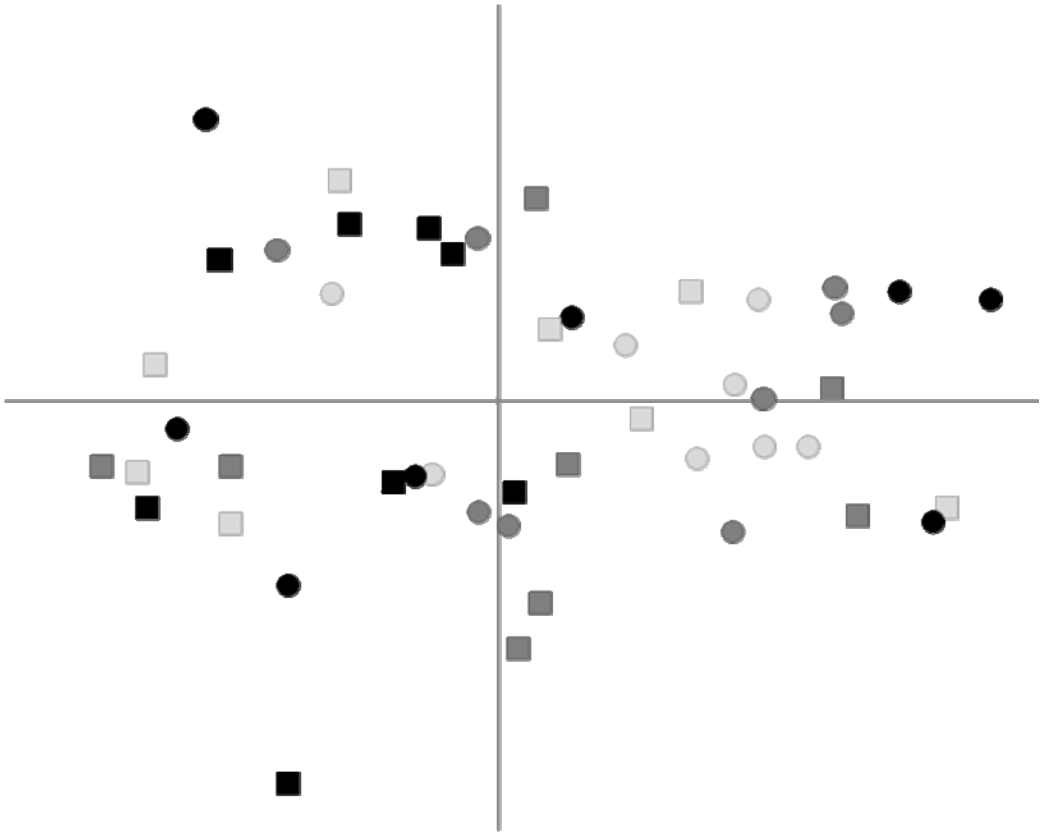

Figure 2. Consensus configuration for the 16 Chinese participants. Dimension 1 (horizontal) explains 27% of the variation in the data; dimension 2 (vertical) explains 25% of the variation. Constructs that load on each dimension are shown, with the number of instances of the construct given in parentheses.

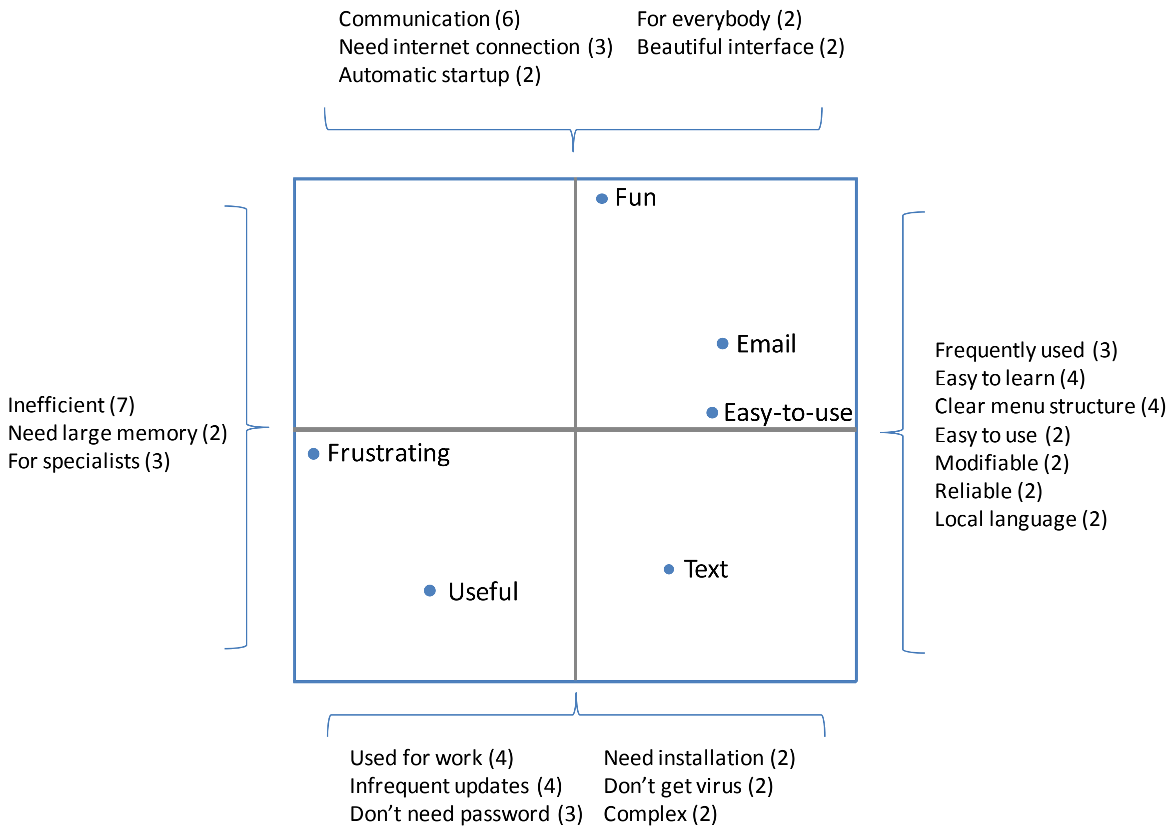

Figure 3. Consensus configuration for the 16 Danish participants. Dimension 1 (horizontal) explains 28% of the variation in the data; dimension 2 (vertical) explains 24% of the variation. Constructs that load on each dimension are shown, with the number of instances of the construct given in parentheses.

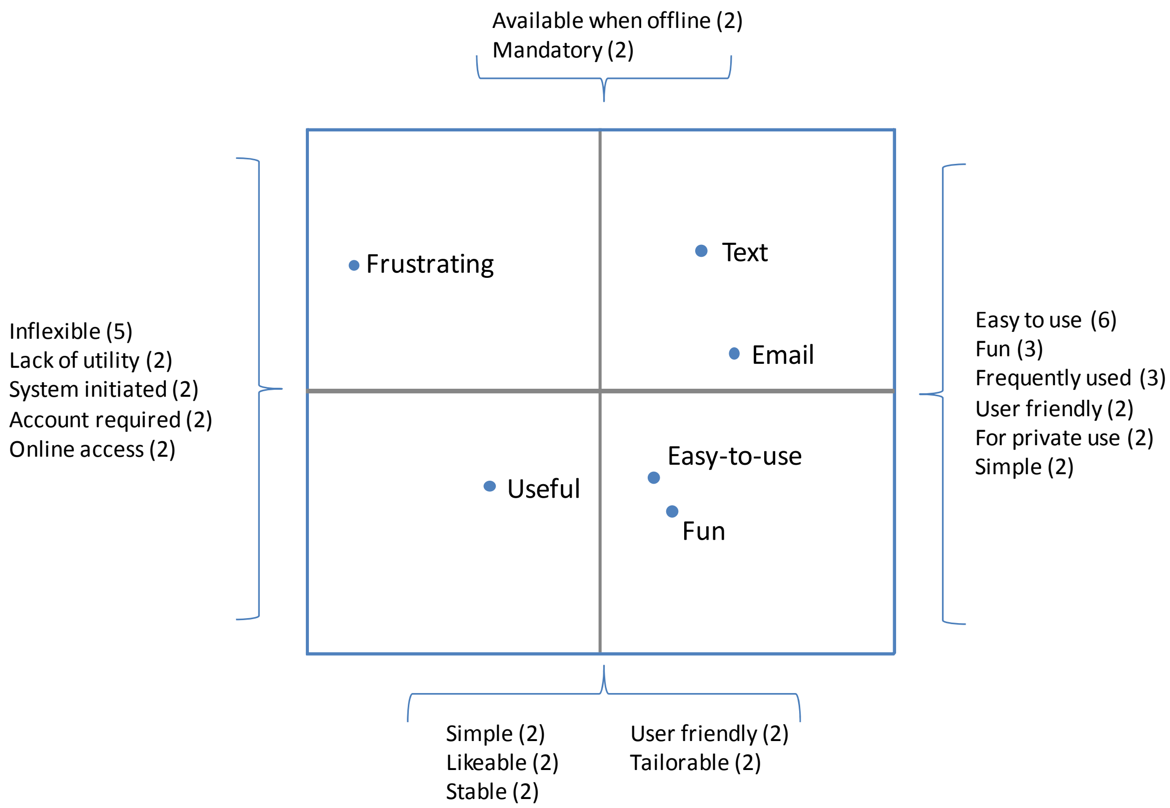

Figure 4. Consensus configuration for the 16 Indian participants. Dimension 1 (horizontal) explains 24% of the variation in the data; dimension 2 (vertical) explains 22% of the variation. Constructs that load on each dimension are shown, with the number of instances of the construct given in parentheses.

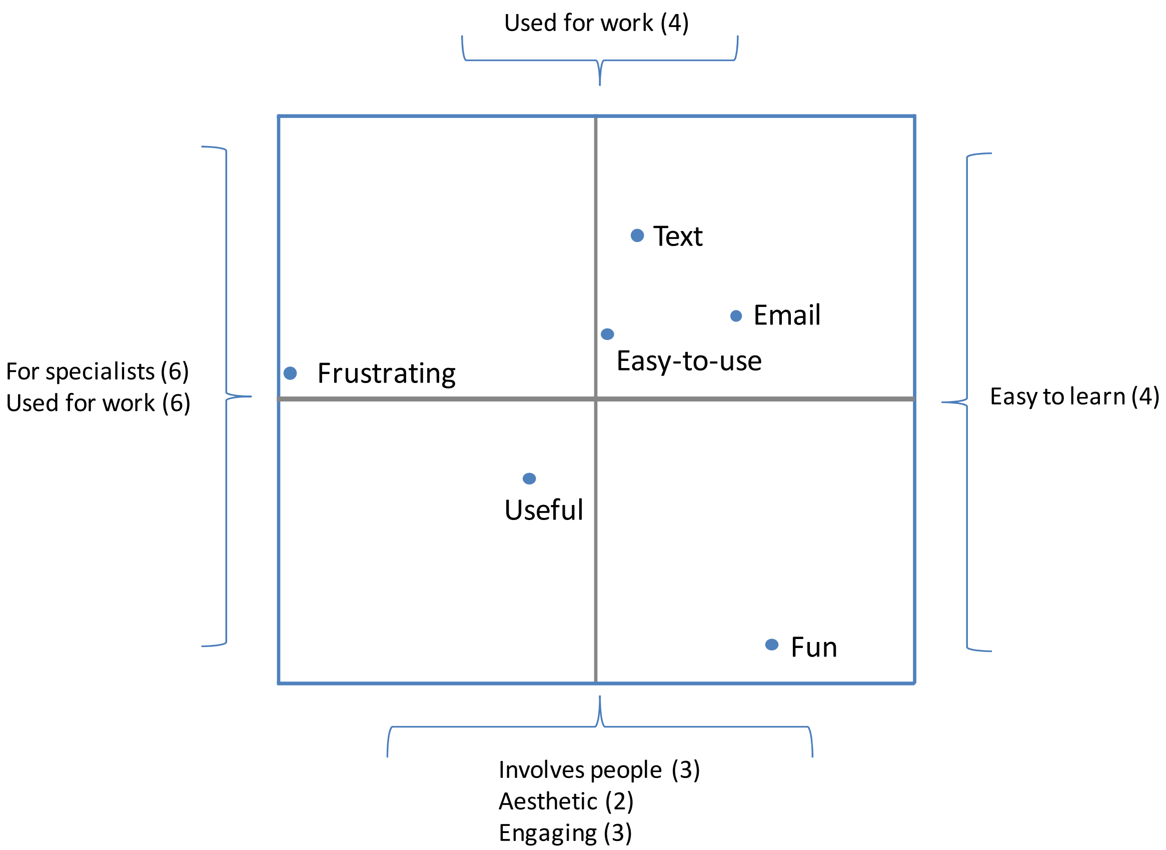

Figure 5. Consensus configuration for the 24 developers. Dimension 1 (horizontal) explains 24% of the variation in the data; dimension 2 (vertical) explains 24% of the variation. Constructs that load on each dimension are shown, with the number of instances of the construct given in parentheses.

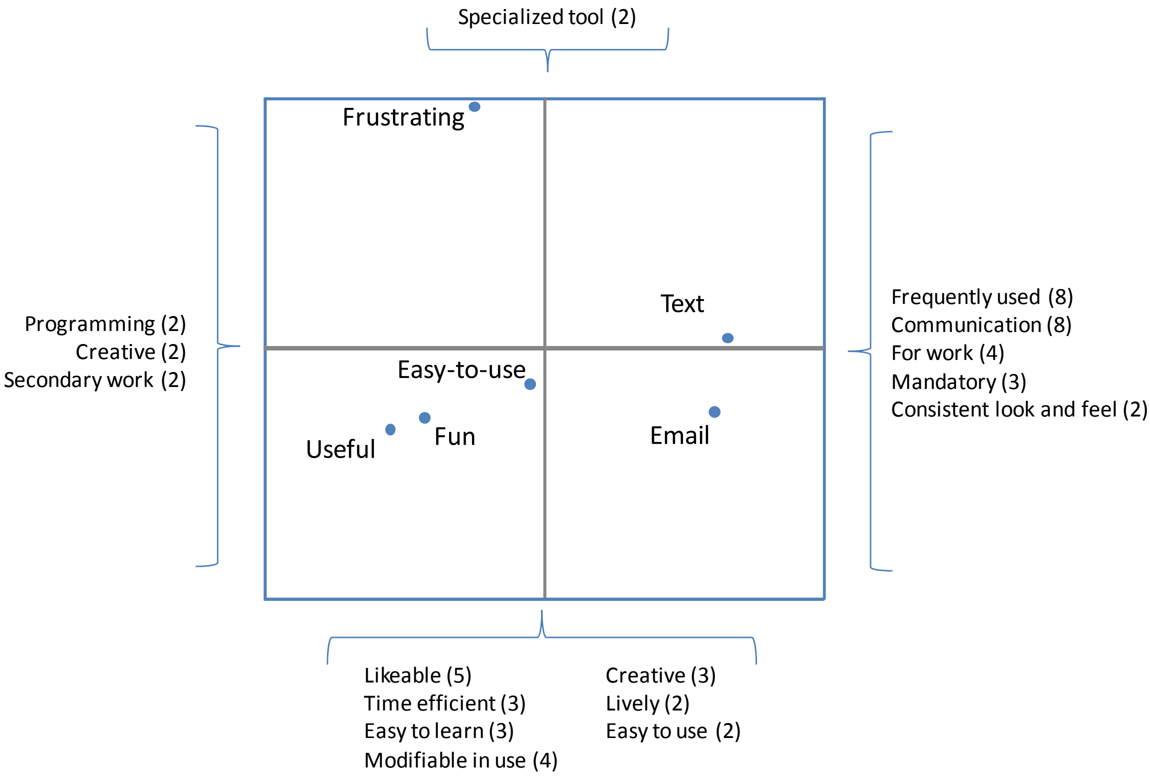

Figure 6. Consensus configuration for the 24 users. Dimension 1 (horizontal) explains 32% of the variation in the data; dimension 2 (vertical) explains 18% of the variation. Constructs that load on each dimension are shown, with the number of instances of the construct given in parentheses.