Non-Universal Usability? A Survey of How Usability is Understood by Chinese and Danish Users

Olaf Frandsen-Thorlacius, Kasper Hornbæk
Department of Computer Science
University of Copenhagen
Njalsgade 128, DK-2300
Copenhagen, Denmark
{olaf,kash}@diku.dk

Morten Hertzum
Computer Science
Roskilde University
Universitetsvej 1, P.O. Box 260
DK-4000 Roskilde
Denmark
mhz@ruc.dk

Torkil Clemmensen
Copenhagen Business School
Department of Informatics
Howitzvej 60
DK-2000 Frederiksberg C
Denmark
tc.inf@cbs.dk

ABSTRACT
Most research assumes that usability is understood similarly by users in different cultures, implying that the notion of usability, its aspects, and their interrelations are constant across cultures. The present study shows that this is not the case for a sample of 412 users from China and Denmark, who differ in how they understand and prioritize different aspects of usability. Chinese users appear to be more concerned with visual appearance, satisfaction, and fun than Danish users; Danish users prioritize effectiveness, efficiency, and lack of frustration higher than Chinese users. The results suggest that culture influences perceptions of usability. We discuss implications for usability research and for usability practice.

Author Keywords
Usability, culture, questionnaire

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation (e.g., HCI): User Interfaces–Evaluation/Methodology.

INTRODUCTION
The concept of usability is fundamental to human-computer interaction (HCI), yet most research assumes that usability is understood similarly across cultures. This assumption is beginning to be questioned by research on cultural usability [e.g., 2,13,20,42,43], which suggests that users may understand usability differently depending on their cultural background. While a number of studies have investigated cultural differences in the graphics, language, object formatting, colors, and layout of web sites and other user interfaces [1,5,30], few studies have addressed whether the concept of usability is perceived differently by users with different cultural backgrounds. This study investigates how Chinese and Danish users understand the concept of usability, its aspects, and their interrelations.

The importance of cultural usability is growing with the increasing numbers of different national and ethnic groups that use information technology on a daily basis. Systems that are marketed worldwide need to be localized to fit cultural specifics [14]; designs that are successful in one culture may not transfer to another [23]; methods of design and evaluation may entail an implicit focus on usability aspects that carry a cultural bias [10]; and definitions of usability such as the ISO 9241 [26] standard may include aspects that are seen as external to usability in some cultural contexts and exclude aspects that are seen as part of usability in others. This way, usability research may make unwarranted claims to universalism because variation in the cultural background of, for example, study participants is not included as an independent variable; often the cultural background of study participants is not even reported.

The number of HCI theories flourish [38], and new ways of measuring usability are being developed [24], but there is little evidence to show how these theories and measures align with attitudes to usability in culturally diverse user groups. In this study we survey two culturally different user populations: Chinese users, which represent an Eastern cultural background, and Danish users, which represent a Western cultural background. The survey questionnaire asked respondents to rate the importance of seven usability aspects and to rank which one in pairs of two usability aspects is the more important to them.

We find significant differences between Chinese and Danish users, and we attempt to characterize these differences, which influence people's understanding of what is or should be the focus of practical usability work as well as usability research. The latter question is relevant because historically important research on the psychology of human-computer interaction [e.g., 6,33,35] is predominantly Western in its origins and may, thus, contain subtle but systematic cultural biases.

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

RELATED WORK

Definitions of usability have evolved with the increasingly diverse range of situations and domains for which systems are being developed. While early definitions conceptualized usability as a narrow, product-oriented quality attribute largely synonymous with ease and simplicity, recent definitions have extended the concept to also include aspects of utility, experience, fun, and culture [e.g., 15,18,43]. As a result, usability has become a diverse concept. In spite of this diversity the ISO 9241 [26] definition of usability has gained widespread acceptance in HCI. According to this definition usability is 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" [26]. This definition is, however, contested outside HCI. For example, another ISO standard [27] provides a usability definition that is directed at software engineering and substantially narrower than the ISO 9241 definition. Moreover, ISO 9241 [26] defines usability in rather abstract terms and this has spawned the development of a layered model, which extends the definition with indicators of usability and means of achieving usability [45].

A common characteristic of most of the abovementioned work is that usability is defined analytically. Comparatively less work attempts to define usability based on users' perception. An exception is McGee et al. [32], who had 46 users rate 64 usability characteristics with respect to how integral the characteristics were to their concept of usability. The results show three clusters of usability characteristic: core usability including consistent, efficient, and easy; secondary usability including effective, controllable, and useful; and tertiary usability including expected and natural. Two additional clusters are separate from the others and therefore appear not to be integral to usability. These two clusters are satisfaction and style. McGee et al. recommend focusing usability activities on core, secondary, and tertiary usability, while excluding satisfaction and style from usability metrics. The study does, however, not consider that users with different cultural backgrounds may perceive usability differently.

The possibility of cultural differences in what constitutes a usable system has become increasingly important as more and more systems become globally available. While a web site becomes available to a worldwide audience the moment it is launched, Hofstede's [21] work on cultural dimensions and Nisbett's [34] work on cultural cognition indicate that a web site designed for users with one cultural background will not be equally usable to users with another cultural background. For example, Marcus [30] describes how power distance, one of Hofstede's cultural dimensions, may affect the prominence that should be given to authorities and symbols on a web site and the directness or discreteness to be applied in using social roles as a basis for differentiated access to information. Sun [43] proposes that cultural usability goes beyond the direct design and use of a system by also involving the social processes concerning how the system is represented, what social identities are associated with it, and what mechanisms regulate its distribution and use. Sun also noted that more research is needed to build an understanding of how culture affects usability. Below, we present six studies that begin to provide such an understanding.

Hertzum et al. [20] performed repertory-grid interviews with 48 people from China, Denmark, and India to elicit their personal usability constructs. For Chinese participants, the most characteristic usability constructs related to security, task types, training, and system issues. In contrast, Danish and to some extent Indian participants more frequently mentioned aspects traditionally associated with usability (e.g., ease of use and intuitiveness). Moreover, a distinction between work and leisure was more widely reported by Indian participants. These results suggest that participants' cultural background influences their usability constructs and that their usability constructs include considerations and distinctions not included in analytic definitions of usability.

O'Keefe et al. [36] had 326 students in United Kingdom, United States, and Hong Kong rate their reactions to the web sites of a European, an American, and an Asian automobile manufacturer. The results show no evidence that the origin of the automobile manufacturer interacts with the user's cultural background. This is a somewhat surprising finding, but it is also difficult to interpret because the students did not view the same version of the web sites but the versions localized for their country. American students found the assessed web sites more relevant and considered them to have higher information content than did British and Hong Kong students. More broadly, the students also differed in their purpose for using the web. American students reported using the web more for information searching and e-commerce compared to Hong Kong students, who reported using it more for social communication and hobby. Based on the same data, Chau et al. [7] report that while Americans dislike web sites with long download times due to heavy graphics, Asians like the graphics, suggesting a difference in preferences for efficiency versus satisfaction.

Evers and Day [16] had 244 students with different cultural backgrounds rate their perception of user interfaces for globally marketed software. Chinese students attached more importance to perceived usefulness in forming an opinion about whether to accept a system, compared to Indonesian students who attached more importance to perceived ease of use. Australian students formed their opinion about whether to accept a system based on neither perceived usefulness nor perceived ease of use.

Choi et al. [9] discerned three cultural dimensions that distinguished how 24 Finnish, Japanese, and Korean interviewees perceived mobile data services. By linking these dimensions (contextuality, uncertainty avoidance, and

individualism/collectivism) with system attributes, Choi et al. provide culture-specific models of what constitute usable mobile data services. For example, the amount of content on a screen should be higher in Korea than Finland and locating information should be simpler in Korea than Japan.

Honold [22] extended the study of cultural aspects of usability from the use of products to include also the process of learning to use them. While Chinese participants preferred a social approach of learning by imitating their friends, German participants preferred an individual approach of learning by doing. This pattern was, however, reversed when participants experienced problems with the products. In face of problems, Chinese participants rarely turned to friends for advice, while German participants were equally likely to ask friends and consult the manual.

Tractinsky [44] had 104 Israeli students rate the beauty and ease of use of 26 layouts of automatic teller machines (ATMs). The study was a replication of a study with Japanese students as participants [29] and found that the correlation between beauty and ease of use was higher among Israeli than Japanese participants. The direction of the difference was contrary to Tractinsky's hypothesis, leading him to the conclusion that current HCI knowledge provides an insufficient basis for accurately predicting how culture influences usability issues.

METHODOLOGY

The aim of the study is to investigate whether cultural background, in this case Chinese or Danish, influences attitudes toward aspects of usability and their interrelation. We investigate this in a questionnaire survey, which allows uniform administration to a large number of respondents.

Respondents

We wanted cultural background to vary among respondents. For this study we focus on one broad cultural difference identified by Nisbett [34], namely that between Westerners (people from Western Europe and US citizens with European origins) and Easterners (people from China and countries heavily influenced by its culture). Of course, numerous other differences between cultures exist, even within the distinction between Westerners and Easterners, but also outside of it [21]. Respondents with an Eastern cultural background came from China, which has high power distance and low individualism. Respondents with a Western cultural background came from Denmark, which has low power distance and high individualism.

Respondents in the study were required to have been born, raised, and currently live in Denmark or China. To ensure that respondents had been raised in accordance with the values of their culture, we also required that both their parents were Danish or Chinese. The number of respondents who fulfilled these criteria (as well as some explained below) was 154 respondents with a Danish background and 258 with a Chinese background.

Themes for the Study

Overall, we hypothesize that cultural background affects the absolute and relative importance users place on aspects of usability, for instance those defined in ISO 9241-11 [26]. More specifically, we wanted to explore four themes that previous work led us to believe would matter for usability. First, we speculate that the visual appearance of a system is given different weight depending on cultural background; Evers and Day [16] presented research that supports such speculation. Second, Hertzum et al. [20] presented preliminary evidence for a number of differences among Chinese and Danish participants, including different emphases on the frustration, fun, and usefulness of systems; we wish to explore whether these differences may be seen also in ratings of perceived usability. Third, the relative importance of effectiveness and ease of use may differ depending on cultural background, as suggested in [16] for Chinese, Indonesians, and Australians. Fourth, we wish to investigate if cultural background carries a preference for efficiency or satisfaction, following the results of Chau et al. [7] on download times versus heavy graphics.

Questionnaire Design

The questionnaire was designed to elicit attitudes about the importance of different aspects of usability (*absolute questions*) as well as to assess attitudes about the relative importance of aspects of usability (*comparison questions*). With this design we aimed to investigate differences in the absolute ratings of aspects of usability, but also wanted to force respondents to prioritize among aspects of usability.

Absolute questions were based on aspects of usability from the literature. Each aspect was measured by a scale consisting of four items (see Table 1). These items were taken from analytic definitions of usability [26] and empirical definitions of usability [20,32]; they also flesh out the constructs in the four themes of the study. Each item was given using a seven-point rating scale with endpoints "not important" (1) and "extremely important" (7). The resulting 28 rating scales comprise the absolute questions.

The comparison questions were created using pairs of items from two aspects of usability. Because all combinations of the seven aspects would result in an overwhelming number of questions, we selected 15 combinations focused on our four themes (see Table 2). The combinations were chosen because they were of particular relevance in relation to related work. Comparison questions were answered on a seven-point rating scale ranging from "only item 1 is important" (1) to "only item 2 is important" (7), with item 1 and item 2 being replaced with the selected combinations of items for the aspects of usability.

To make respondents respond on the basis of their concrete experiences, questions were asked in relation to two systems: respondents' e-mail program and their word processor. The reasons for including two systems are that previous research has found differences in the emphasis placed on usability aspects for different systems [20], that

we get more reliable data because questions are answered twice, and that answers to questionnaires are more reliable when given in relation to a concrete experience [e.g., 41]. The questionnaire therefore comprises a total of 86 questions about usability (two times 28 absolute questions, plus two times 15 comparison questions).

The questionnaire was developed in an English master version, summarized in tables 1 and 2. The master version was translated into Danish and Chinese. The questionnaire was administered in the native language of respondents for two reasons: to give respondents a better understanding of the questionnaire and to ensure that no respondents were excluded due to language difficulties. To check the translation, questionnaires were translated back to English and checked against the master version; such back-translation is a common way of ensuring validity in cross-cultural research [4].

Consider the following as an example of the translation of the questions into Danish and Chinese. In the English master version one question read "How important is it to you that your text editing program is: interesting to look at", allowing seven answers from "not important" to "extremely important". This was translated into Danish as "hvor vigtigt er det for dig at dit tekstbehandlingsprogram er interessant at se på" (with answer possibilities from "ikke vigtigt" to "ekstremt vigtigt") and into Chinese as "对您来说，这个对您的电子邮件的重要性为: 看上去令 人有兴趣的" (with answer possibilities from "不重要" to "极其重要").

The questionnaire was pilot tested in three steps, first using a think-aloud test, then by having 8 subjects complete it in English, then by having three persons complete the Danish version and three persons complete the Chinese version. The modifications resulting from the pilot test have been incorporated in the questions presented in tables 1 and 2.

The literature on cultural differences [e.g., 21,34] led us to think that the importance of support for social interaction will differ among cultural backgrounds; a recent survey of cross-cultural effects in computing also called for a better

Scale	Alpha	Item wording
Effectiveness	.866	Useful, Productive, Effective, Increases performance
Ease of use	.896	Easy to use, Clear and Understandable, Simple, Quick to learn
Visual appearance	.916	Beautiful, Lively appearance, Inspiring to look at, Interesting to look at
Efficiency	.897	Fast, Efficient, Swift, Starts quickly
Satisfaction	.856	Feels friendly, Comfortable, Interesting to use, Inspiring to use
Fun	.917	Fun, Likeable, Enjoyable, Amusing
Non-frustration	.894	Non-frustrating, Non-annoying, Pleasant, Non irritating

Table 1. Scales in the questionnaire, their inter-item reliability (Cronbach's alpha [12]), and the wording of the items (in the absolute questions).

Comparison	Alpha	Item wording
Ease of use vs. effectiveness	.663	Simple vs. Productive, Easy to use vs. Useful, Improves performance vs. Clear and understandable.
Visual appearance vs. efficiency	.787	Beautiful vs. Fast, Inspiring to look at vs. Efficient, Starts quickly vs. Interesting to look at
Satisfaction vs. efficiency	.801	Efficient vs. Interesting to Use, Inspiring to use vs. Starts quickly, Beautiful vs. Fast.
Fun vs. ease of use	.833	Fun to use vs. Simple, Clear and understandable vs. Likable, Fun to use vs. Easy to Use
Effectiveness vs. non frustrating	.773	Improves performance vs. Non-frustrating, Productive vs. Non-annoying, Non-frustrating vs. useful

Table 2. Comparison questions, their inter-item reliability (Cronbach's alpha [12]), and the wording of the items (in the comparison questions).

treatment of social phenomena [37]. Thus, we asked respondents two questions for their e-mail program and word processor (using the same scales as for the absolute questions: "Supports social activities" and "Supports communication other than E-Mail". These questions will help understand whether any differences between e-mail program and word processor are due to their support of social activity and whether the importance of such activity depends on respondents' cultural background.

Reliability

We tested the inter-item reliability of questions for each scale using Cronbach's alpha [12]. Alpha indicates the extent to which questions correlate with each other. A scale is typically considered reliable if its alpha value is above a threshold of .7. As may be seen in Table 1, the seven scales of absolute questions had alphas ranging from .856 to .917. The five comparison scales had lower reliability, with alphas ranging from .663 to .833. We kept the comparison scale "ease of use versus effectiveness" even though it was slightly unreliable because no obvious part of the scale could be removed to improve its reliability.

This analysis confirms that the scales and the comparison questions are reliable indicators of the aspects of usability and the comparisons of aspects of usability we chose to investigate. Later in the paper we conduct factor analysis to investigate how items and aspects relate to each other.

Administration and Procedure

The questionnaire was distributed online using surveymonkey.com. Potential respondents were contacted using e-mail and by posting on web forums. Respondents were offered to enter a draw for two music player gifts.

The questions described in Table 1 and Table 2 were preceded by a series of background questions. The background questions concerned respondents' age, gender, and educational level; their use of e-mail programs (five options, Never to Daily); their use of word processors (again, Never to daily); and their cultural background

(country of birth, where they were raised, where they live now, and their parents' nationality).

Data Processing and Analysis

From an initial pool of 528 responses to the questionnaire, we removed responses where questions about cultural background were not all answered, and where answers to the questions on cultural background were not similar (e.g., born in China but raised in the US). We also excluded one response with less than 25% of the questions answered, leaving 412 responses for further analysis.

A total of 2.7% missing values were replaced with the value of series means to allow for factor analysis and full statistical analysis. We analyzed data using multivariate analysis of variance (MANOVA), with cultural background as the independent variable. Our dependent variables were either the seven aspects of usability or the 86 individual items. Age, gender, and educational level were used as covariates. Finally, we used factor analysis to understand better the match between our a priori scales and the empirical structure of answers.

RESULTS

Demographics and Software Use

Answers from 412 respondents, 154 Danish and 258 Chinese, were used for the analysis. The average age of respondents was 26.4 years (*SD* = 5.92). On average, Danish respondents were 3.82 years older than Chinese respondents. An equal number of male and female respondents participated. Chinese respondents encompassed 53% female and 47% male, while Danish respondents encompassed 45% female and 55% male. The educational level of both groups was high, 95% had attended college.

Respondents in both cultural groups used e-mail frequently (Danish: 95% daily; Chinese: 89% daily) and used a word processor at least on a weekly basis (Danish: 95% weekly, 55% daily; Chinese: 92% weekly, 67% daily).

Absolute question	Chinese		Danes		F(1,410) =
	M	SD	M	SD	
Effectiveness	4.80	1.09	5.66	0.76	75.15, p<.001
Ease of use	5.10	1.09	4.96	1.17	1.98, p>.2
Visual appearance	4.20	0.94	2.50	1.01	301.39, p<.001
Efficiency	5.36	1.10	5.66	0.89	8.56, p<.05
Satisfaction	4.29	1.02	3.39	0.93	77.67, p<.001
Fun	3.74	102	1.97	0.82	335.43, p<.001
Non-frustration	4.46	1.10	5.25	0.87	57.19, p<.001

Table 3. Means, standard deviations, and tests of significant difference (ANOVAs) for absolute questions.

Absolute Questions

Figure 1 shows the averages of the scales created from respondents' answers to the absolute questions. As may be seen from the figure, cultural differences exist in the answers to the questions; this is confirmed by an overall multivariate analysis of variance with the seven scales as dependent variables, $F(7, 404) = 120.31$, $p < .001$. Following Cohen's terminology on effect size [11], the effect size of the difference between cultural backgrounds is moderate. Below we analyze each scale in turn (see Table 3 for statistics).

Effectiveness was significantly more important to Danish respondents, rated about 18% higher than for Chinese respondents. The difference in ratings is consistent across items, with median scores of the eight questions differing by one step in seven cases and being equal in one.

For *ease of use*, we find no significant difference between Chinese and Danish respondents. Median scores of the eight questions on ease of use are identical in five cases, higher for Danish respondents in two cases, and higher for Chinese respondents in one case. Somewhat surprisingly, we find a strong effect of the gender covariate for ease of use. Female respondents attach more importance to ease of use ($M = 5.34$, $SD = 1.05$) compared to male respondents ($M = 4.76$, $SD = 1.12$). This gender effect does not interact with cultural background, nor does it seem to otherwise

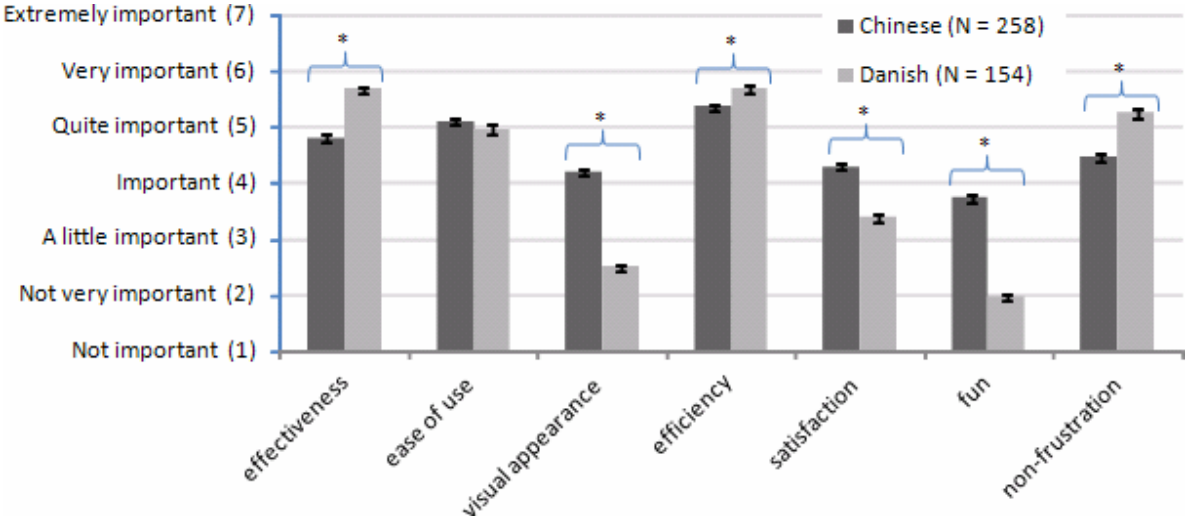

Figure 1. Average responses to the seven scales, each consisting of four questions. Six scales, marked with asterisks, show a significant difference between Chinese and Danish respondents. Error bars show the standard error of the mean.

mask an effect of culture.

Visual appearance differs markedly between Chinese and Danish respondents. Chinese respondents consider visual appearance important, while Danish respondents consider it to be between not very important and a little important (40% lower than Chinese respondents). This difference is significant and has a medium effect size ($eta^2 = .42$).

We find a significant difference between Danish and Chinese respondents with respect to *efficiency*. However, this effect is very small (5-6%, $eta^2 = .02$). We find a significant and stronger correlation between efficiency scores and use frequency (Spearman correlation coefficient $r = .149$), suggesting that the more experience respondents had with e-mail program or word processing, the more importance they placed on efficiency. In concrete terms, daily users of e-mail programs or word processors have 11% higher scores on the efficiency scale compared to non-daily users, or about half a step on the rating scale.

We find a significant difference between cultural backgrounds with respect to *satisfaction*. Chinese respondents found satisfaction important, while Danish respondents only found it to be of little importance. This is a significant difference, with ratings of Chinese respondents 27% higher than Danish respondents. Interestingly, answers to the item on comfort differ from answers to the other three items making up the satisfaction scale. For the comfort item Danish respondents have a higher mean rating ($M = 5.45$, $SD = 1.11$) than Chinese respondents ($M = 5.23$, $SD = 1.20$). The six other questions show the opposite pattern. We return to this observation in the factor analysis.

Fun is considered significantly more important by Chinese respondents compared to Danish respondents. Danish respondents attached only little importance to fun, while Chinese respondents were about 1.8 scale steps higher on the seven-step rating scale. This difference is the largest among the seven aspects of usability: in terms of median scores for the eight questions, Chinese respondents answered one scale step higher for four questions, two scale steps higher for two questions, and three scale steps higher for two questions (the enjoyability item for both e-mail program and word processing).

Non-frustration is of significantly higher importance to Danish respondents compared to Chinese respondents. Again this may be illustrated with the difference in median values, which in four cases is two, in one case is 1.5, and in three cases is one. The differences of two scale steps are for the items non-frustration and non-annoyance (for both e-mail program and word processing).

Comparisons between Usability Aspects

Figure 2 summarizes the respondents' comparisons between pairs of usability aspects. Using MANOVA, we find a significant overall difference between cultural background using the five comparison scales as dependent variables, $F(5, 406) = 72.03$, $p < .001$. The size of the effect of cultural background is medium ($eta^2 = 0.47$). Next we discuss each comparison question (Table 4 gives individual statistics); overall they corroborate the differences illustrated in Figure 1.

The comparison between *ease of use* and *effectiveness* shows only a small, though significant, difference between Chinese and Danish respondents. Consistent with the absolute scales, Danish respondents attached more importance to effectiveness compared to Chinese respondents. As suggested by the absolute questions, the

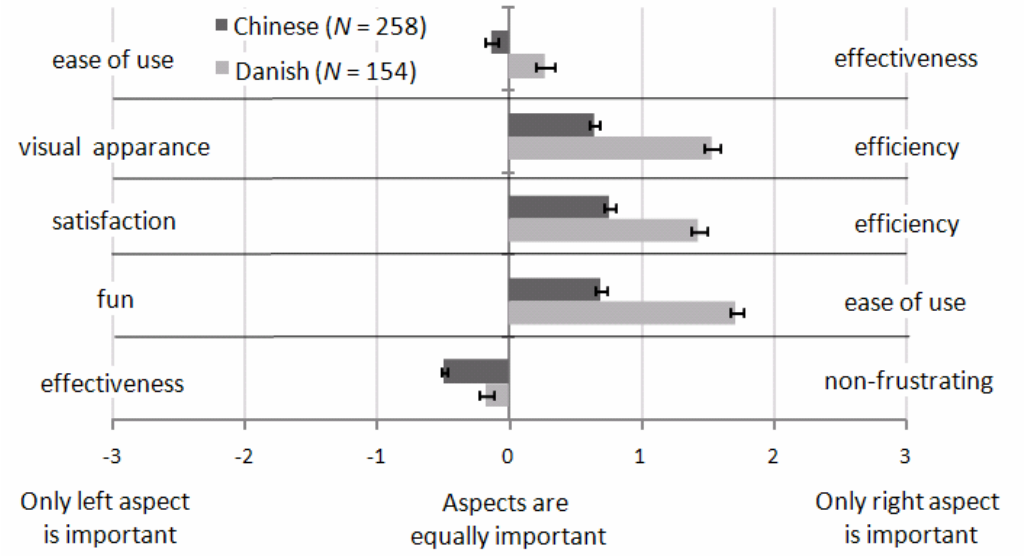

Figure 2. Respondents' comparisons between aspects of usability. Each row shows the relative importance of the two aspects of usability indicated at the ends of the row. Error bars show the standard error of the mean. All five comparison scales are significantly affected by cultural background.

Comparison question	Chinese		Danes		F(1,410) =
	M	SD	M	SD	
Ease of use vs effectiveness	-0.16	0.91	0.31	1.13	17.61, p<.001
Visual appearance vs efficiency	0.74	0.74	1.78	0.89	159.11, p<.001
Satisfication vs efficiency	0.88	0.74	1.65	0.89	89.83, p<.001
Fun vs ease of use	0.80	0.77	1.98	0.79	226.35, p<.001
Effectiveness vs non-frustration	-0.58	0.87	-0.21	0.86	17.61, p<.001

Table 4. Means, standard deviations, and ANOVAs for comparison questions. Negative values indicate a preference for the first aspect in the comparison question.

effect size is small ($eta^2 = .04$).

The comparison of *visual appearance* versus *efficiency* is significantly affected by cultural background. Chinese respondents considered these two aspects of usability about equally important with a tendency to consider efficiency "slightly more important"; Danish respondents differs by about one point on the rating scale and were much closer to saying that efficiency "is the most important".

Efficiency is also significantly preferred over *satisfaction*. Reflecting the absolute questions, Chinese respondents assessed the relative importance of efficiency and satisfaction as more equal than Danish respondents.

In the comparison of *ease of use* versus *fun*, we again find a significant influence of cultural background. Among Chinese respondents, ease of use was seen as the only important property (rating 3) in only 4% of the answers; among Danish respondents, ease of use was considered the only important property in 48% of the answers. This confirms that fun is understood and prioritized very differently depending on cultural background.

For *effectiveness* versus *non-frustration*, respondents preferred effectiveness. Though the difference between the ratings of Chinese and Danish respondents was significant, it was small (about a quarter of a scale step).

Differences between Software

Using MANOVA on the means of the seven usability aspects for the two kinds of software, we find a difference between e-mail program and word processing, $F(7, 404) = 9.61$, $p < .001$. Using Bonferroni corrected post-hoc tests, we find only one significant aspect involved in this effect: effectiveness is mediated by both a significant effect of type of software, $F(1, 410) = 47.84$, $p < .001$, and by an interaction between software and cultural background, $F(1, 410) = 15.29$, $p < .001$. Closer analysis of this effect shows that higher importance was placed on effectiveness by Danish respondents for word processing ($M = 5.90$; $SD = .85$) than for e-mail program ($M = 5.42$; $SD = .85$); Chinese respondents assessed these similarly (e-mail program: $M = 4.73$; $SD = 1.14$; word processing: $M = 4.86$; $SD = 1.22$).

Apart from this effect, the average ratings on the seven usability aspects were similar for e-mail program and word processing, with average differences between the two kinds of software being 2% (ranging from .7% to 5%).

Correlations among Scales and Factor Analysis

Table 5 shows the pair-wise correlations between the seven aspects of usability. The table shows that in general the correlations among variables are medium to large, with some aspects (e.g., satisfaction) being related significantly to all other aspects. Visual appearance and fun appear to be distinct from the other aspects of usability, except satisfaction.

To investigate further the relations among the respondents' answers to the questionnaire we performed a factor analysis [17] on the 56 absolute questions (using varimax rotation). The analysis yielded eight components with eigenvalues above 1, explaining 64% of the variance. Kaiser-Meyer-Olkin measure of sampling adequacy was .946, suggesting marvelous common variance and factorability of the data [28]. Inspection of the scree plot and of unique loadings suggested six components should be retained, explaining 60.6% of the variance. One item of the Non-frustration scale failed to load above .5 on any component (Pleasant). As items ended up in the same factor, independently of being asked of the e-mail program or the word processor; we present results for items below.

Table 6 shows the factors identified. The first point from the factor analysis is that the two components with the highest loadings correspond well to a subjective and an objective dimension of usability. The *experience* component (explaining 27% of the variance) has loadings above .5 on all items on visual appearance and fun, and on three of four items on satisfaction. The other highly loaded component concerned *performance* (explaining 19% of the variance), and concerned items on efficiency and the item about usefulness from the effectiveness scale. The remainder of the components account for less variance.

The components may be used to generate factor scores, which may in turn be tested for differences across cultural

	Effectiveness	Ease of use	Visual appearance	Efficiency	Satisfaction	Fun
Effectiveness	1					
Ease of use	.42*	1				
Visual appearance	.00	.37*	1			
Efficiency	.67*	.63*	.14	1		
Satisfaction	.26*	.49*	.78*	.30*	1	
Fun	-.03	.25*	.84*	.06	.76*	1
Non-frustration	.66*	.37*	-.09	.46*	.22*	-.12

Table 5. Correlations among aspects of usability ($N = 412$ respondents). Significant correlations using Bonferroni adjustments are flagged with asterisks.

backgrounds [17]. This type of analysis confirms the earlier reported differences between cultural backgrounds, $F(8, 403) = 112.55$, $p < .001$. The empirically derived factors show results similar to the scales defined a priori, using usability aspects from the literature. We find significant differences across cultural backgrounds in the experience component ($F[1, 410] = 275.43$, $p < .001$), the efficiency component ($F[1, 410] = 4.47$, $p < .05$), the lack-of-frustration component ($F[1, 410] = 132.91$, $p < .001$), and the effectiveness component ($F[1, 410] = 9.29$, $p < .01$).

The factor analysis also reveals a couple of issues that did not behave as expected. The part of the satisfaction scale that asked about comfort ended in its own component, explaining 2% of the variance. As discussed above, this question appears to be unrelated to the other items making up the satisfaction aspect. Further, the questions about effectiveness were not related as expected: the item on usefulness was related to efficiency, which was unexpected given analytic definitions of usability (e.g., ISO 9241-11).

The Role of Social Activity

The questions on social activity differed significantly across cultural backgrounds, $F(1, 410) = 73.34$, $p < .001$. Chinese respondents placed more importance on social activity ($M = 3.39$; $SD = 1.21$) than Danish respondents ($M = 2.39$; $SD = 1.02$). While Chinese respondents appeared to value support for social activity, they still rated other aspects of usability higher. The usability aspect with which the questions on social activities correlated most strongly was fun ($r = .60$, $p < .001$), but social activity also correlated significantly with three of the six other aspects of usability (rs in the range .21 to .53, all ps < .05)

DISCUSSION

Our study shows that the answers to questions about the absolute and relative importance of aspects of usability differ across cultural backgrounds. Chinese respondents place more emphasis on visual appearance, satisfaction, and fun than Danish respondents, who in turn place more importance on effectiveness and the lack of frustration. These effects are medium-sized and hold across absolute and comparison questions, except for efficiency, which only shows a small effect of cultural background. The structure of answers suggests that factors on experience and performance capture much of the variation in the responses to the survey, and that these two factors are considered differentially important across cultural backgrounds.

Component label	Variance	Items (with loadings >.5)
Experience*	27.4%	Visual appearance (all items) Fun (all items) Satisfaction (3/4 items)
Performance*	19%	Efficiency (all items) Effectiveness ("useful" item)
Lack of frustration*	5.5%	Non-frustration (3/4 items)
Ease of use	3.4%	Ease of use (all items)
Effectiveness*	2.8%	Effectiveness (2/4 items)
Comfort	2.3%	Satisfaction ("comfortable" item)

Table 6. Factor Analysis of Absolute Questions. Asterisks show significant differences between cultural backgrounds when comparing their factor scores.

In relation to the themes put forward earlier in the paper, the results confirm that the perceived importance of visual appearance differs between cultural backgrounds, generalizing the conclusions of [7]. Similarly to Hertzum et al. [20], we find differences among Chinese and Danish respondents with respect to frustration and fun; differences in the perceived importance of usefulness were less clear. With respect to the relative importance of effectiveness and ease of use, our results suggest some differences owing to cultural background; particularly, Danish respondents prioritize effectiveness to a larger extent than Chinese respondents. Effectiveness and efficiency seem to define one important factor along which respondents differ in their answer. We found support for the speculation that the relative importance of efficiency and satisfaction may differ across cultural backgrounds, with Danish respondents prioritizing efficiency compared to Chinese respondents. In contrast to [7], the absolute preferences, however, were in favor of efficiency. Finally, the importance of social activity appears related to respondents' understanding of usability and is affected by cultural background.

Impact on Usability Research and Future Work

The results have at least four implications for usability research. First, while widespread models of usability may capture the aspects of usability, they generally fail to recognize cultural variation in the importance of those aspects. For instance, the ISO 9241-11 [26] mentions effectiveness as one of three core aspects of usability. Our results suggest that its importance differs with cultural background. Though work on the Technology Acceptance Model (TAM) has suggested culture as a mediating factor in technology acceptance [16,40], we believe that our study provides the first quantitative evidence of the influence of cultural background on usability. In contrast to TAM research, our work concerns aspects of usability in a broad sense and on their perceived value and relative importance.

Second, empirical results on the relation between aspects of usability [25,31,39] consistently fail to account for and discuss the influence of cultural background. For instance, Hornbæk and Law [25] did not discuss any moderating effects of cultural background in their meta-analysis of correlations among usability aspects. While our results concern perceived usability and are not obtained after a recent use situation common to respondents, we speculate that other measures of usability might also be affected by cultural background. Certainly, we expect them to influence ratings of usability measured by, for instance, standardized questionnaires such as QUIS [8].

Third, our study supports the emerging focus on validating and extending definitions of usability based on empirical studies [e.g., 32]. In contrast to the results of McGee et al. [32], we do not conclude that satisfaction qualities "are not

perceived as integral to usability by users" (p. 908). For Danish respondents, this conclusion seems reasonable, as they rate satisfaction lowly (see Figure 1); for Chinese respondents, however, it seems unwarranted as they rate satisfaction qualities comparably to, say, effectiveness. We also find that among the satisfaction-related aspects (i.e., satisfaction, non-frustration, fun), fun appears to be the least important to users. This could be seen as supporting the ISO 9241 definition of usability, which almost excludes fun from the satisfaction aspect, and to support Hassenzahl et al. [19] and others in arguing that fun is beyond usability. A puzzling issue is that the scale on absence of frustration appears to be different from satisfaction in that it has only a low correlation with it (see Table 5).

Fourth, our questions on social activities suggest other relevant concerns for models of usability. Does support for social activities need to be considered in relation to models of usability? Why are social activities correlated most strongly with fun and not, as one might expect, with effectiveness or usefulness? Our study does not answer these questions; we merely propose them for future work.

Impact on Usability Practice

An obvious question is how our results should affect practical usability testing and user-centered design. The key message is that findings on perceived usability are not transferable across all kinds of cultural background. Increasingly, usability work is international [e.g., 3] and our findings caution against universalism about the importance of different aspects of usability. We are also skeptic about the possibility of using and comparing results from standardized satisfaction questionnaires across different cultural backgrounds. Also, usability professionals doing usability testing with users from multiple cultural backgrounds should be aware of potential differences in their perception of usability. Finally, our results suggest that practitioners should give priority to different aspects of usability in order to develop systems that will be perceived as usable by users with different cultural backgrounds.

Limitations and Possible Objections

Our study has several limitations; we also want to briefly discuss a couple of possible objections to our methodology. First of all, the questions to respondents purposefully used programs that could be used for both private and work purposes. However, the choice of software may bias our results. For instance, would results be different if respondents had been answering the same questions in relation to, say, games, chat systems, or e-banking? The analysis of the data may also be extended using structural equation modeling, and more formal comparisons of how the factor structure of responses differs between Chinese and Danish respondents. In relation to the methodology of the study it may, in particular, be objected that it depends crucially on the translation process which, despite our use of back translation, could be validated in other ways. Also, the comparison questions may have an inherent cultural bias in that Chinese respondents may prefer the middle of the comparison scale [34].

We acknowledge that the results of the present survey need replication in real-life use of software. Future work should aim at testing whether the differences found here translate into differences in participants' responses when asked after having used the same system, for instance in a laboratory. Finally, it would be interesting to design software scoring high or low on the dimensions above and investigate whether that affects overall preference and adoption.

CONCLUSION

Most research in human-computer interaction seems to assume that usability is understood similarly across cultures. We challenged this assumption by investigating the effects of cultural background (viz., Chinese and Danish) on seven aspects of usability using a questionnaire. Respondents were asked to assess the importance of these aspects for their e-mail program and word processor.

The results show that Chinese respondents place more importance on visual appearance, satisfaction, and fun than Danish respondents. Conversely, Danish respondents report effectiveness, efficiency, and the absence of frustration as more important than do Chinese respondents. Ease of use is considered equally important by Chinese and Danish respondents. Among the satisfaction-related aspects of usability both Chinese and Danish respondents consider non-frustration the most important and fun the least important. The importance attached to social activity is also affected by cultural background, suggesting cultural differences beyond the seven aspects of usability.

Our findings suggest that perceived usability, for instance as measured in satisfaction questionnaires, is affected by the cultural background of participants. Caution is also needed in interpreting the results of cross-cultural usability tests. We argue that usability research needs to look more into cultural background as a moderator of preferences and of the relation between usability aspects and preferences.

ACKNOWLEDGEMENTS

We wish to thank Qingxin Shi, Duo Duo Zhuang, and Naja Rosenkjær Villien for assistance with translation, and our respondents for their participation. This study was co-funded by the Danish Council for Independent Research through its support of the Cultural Usability project.

REFERENCES

1. Aykin, N., Quaet-Faslem, P. H., & Milewski, A. E., Cultural Ergonomics, in Salvendy, G. (ed.) *Handbook of Human Factors and Ergonomics,* Wiley, 2006, 177-190.
2. Barber, W. & Badre, A. Culturability: The Merging of Culture and Usability, *Proc. of the Fourth Conference on Human Factors and the Web* (1998).
3. Bojko, A., Lew, G. S., & Schumacher, R. M. Overcoming the Challenges of Multinational Testing, ACM *interactions*, *12*, 6 (2005), 28-30.

4. Brislin, R. W. Back-Translation for Cross-Cultural Research, *Journal of Cross-Cultural Psychology*, *1*, 3 (1970), 185-216.
5. Callahan, E., Interface Design and Culture, in Cronin, B. (ed.) *Annual Review of Information Science and Technology, Vol. 39,* Information Today, 2005, 257-310.
6. Card, S., Moran, T., & Newell, A. *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum, Hillsdale NJ, 1983.
7. Chau, P. Y. K., Cole, M., Massey, A., Montoya-Weiss, M., & O'Keefe, R. M. Cultural Differences in the Online Behavior of Consumers, *Communications of the ACM*, *45*, 10 (2002), 138-143.
8. Chin, J. P., Diehl, V. A., & Norman, K. L. Development of an Instrument for Measuring User Satisfaction of the Human-Computer Interface, *Proc. CHI '88*, ACM Press (1988), 213-218.
9. Choi, B., Lee, I., & Kim, J. 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 (2006), 171-206.
10. Clemmensen, T., Hertzum, M., Hornbæk, K., Shi, Q., & Yammiyavar, P. Cultural Cognition in the Thinking-Aloud Method for Usability Evaluation, *Proc. ICIS 2008*, AIS (2008)
11. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, Academic Press, New York, 1969.
12. Cronbach, L. J. Coefficient Alpha and the Internal Structure of Tests, *Psychometrika*, *16* (1951), 297-333.
13. Day, D. Ed. Shared Values and Shared Interfaces [Special Issue], *Interacting with Computers*, *9*, 3 (1998).
14. del Galdo, E. M. & Nielsen, J. *International User Interfaces*, Wiley, New York, 1996.
15. Elliott, M. & Kling, R. Organizational Usability of Digital Libraries: Case Study of Legal Research in Civil and Criminal Courts, *Journal of the American Society for Information Science*, *48*, 11 (1997), 1023-1035.
16. Evers, V. & Day, D. The Role of Culture in Interface Acceptance, *Proc. Interact'97*, Chapman and Hall (1997), 260-267.
17. Gorsuch, R. L. *Factor Analysis*, Lawrence Erlbaum, Hillsdale, NJ, 1983.
18. Hassenzahl, M., Beu, A., & Burmester, M. Engineering Joy, *IEEE Software*, *18*, 1 (2001), 70-76.
19. Hassenzahl, M., Platz, A., Burmester, M., & Lehner, K. Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal, *Proc. CHI 2000*, ACM Press (2000), 201-208.
20. Hertzum, M., Clemmensen, T., Hornbæk, K., Kumar, J., Shi, Q., & Yammiyavar, P. Usability Constructs: A Cross-Cultural Study of How Users and Developers Experience Their Use of Information Systems, *Proc. HCI International, LNCS 4559*, Springer Verlag (2007), 317-326.
21. Hofstede, G. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, Sage, Thousand Oaks, CA, 2001.
22. Honold, P. Learning How to Use a Cellular Phone: Comparison Between German and Chinese Users, *Technical Communication*, *46*, 2 (1999), 196-205.
23. Honold, P. Culture and Context: An Empirical Study for the Development of a Framework for the Elicitation of Cultural Influence in Product Usage, *International Journal of Human-Computer Interaction*, *12*, 3&4 (2000), 327-345.
24. Hornbæk, K. Current Practice in Measuring Usability: Challenges to Usability Studies and Research, *International Journal of Human-Computer Studies*, *64*, 2 (2006), 79-102.
25. Hornbæk, K. & Law, E. L. C. Meta-Analysis of Correlations Among Usability Measures, *Proc. CHI 2007* (2007), 617-626.
26. ISO 9241-11 Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs)-Part 11: Guidance on Usability, International Standard Organization (1998).
27. ISO/IEC 9126 Software Engineering - Product Quality - Part 1: Quality Model, International Standard Organization (2001).
28. Kaiser, H. F. An Index of Factor Simplicity, *Psychometrika*, *39* (1974), 31-36.
29. Kurosu, M. & Kashimura, K. Determinants of Apparent Usability, *Proc. IEEE International Conference on Systems, Man and Cybernetics*, IEEE Press (1995), 1509-1513.
30. Marcus, A., Global/Intercultural User Interface Design, in Sears, A. & Jacko, J. (eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications,* Erlbaum, 2008, 355-390.
31. McGee, M. Master Usability Scaling: Magnitude Estimation and Master Scaling Applied to Usability Measurement, *Proc. CHI 2004*, ACM Press (2004), 335-342.
32. McGee, M., Rich, A., & Dumas, J. Understanding the Usability Construct: User-Perceived Usability, *Proc. HFES'2004*, HFES (2004), 907-911.
33. Nielsen, J. *Usability Engineering*, Academic Press, San Diego CA, 1993.
34. Nisbett, R. E. *The Geography of Thought: How Asians and Westerners Think Differently - and Why*, Nicholas Brealey, London, 2003.
35. Norman, D. A. & Draper, S. W. *User Centered System Design*, Lawrence Erlbaum, Hillsdale, NJ, 1986.
36. O'Keefe, R. M., Cole, M., Chau, P. Y. K., Massey, A., Montoya-Weiss, M., & Perry, M. From the User Interface to the Consumer Interface: Results From a Global Experiment, *International Journal of Human-Computer Studies*, *53*, 4 (2000), 611-628.
37. Rau, P.-L. P., Gao, Q., & Liang, S.-F. M. Good Computing for Everyone - How on Earth? Cultural Aspects, *Behaviour & Information Technology*, *27*, 4 (2008), 287-292.
38. Rogers, Y., New Theoretical Approaches for Human-Computer Interaction, in Cronin, B. (ed.) *Annual Review of Information Science and Technology, Vol. 38,* Information Today, 2004, 87-144.
39. Sauro, J. & Kindlund, E. A Method to Standardize Usability Metrics into a Single Score, *Proc. CHI 2005*, ACM Press (2005), 401-409.
40. Schepers, J. & Wetzels, M. A Meta-Analysis of the Technology Acceptance Model: Investigating Subjective Norm and Moderation Effects, *Information & Management*, *44*, 1 (2007), 90-103.
41. Schwarz, N., Groves, R., & Schuman, H., Survey Methods, in Gilbert, D., Fiske, S., & Lindzey, G. (eds.) *Handbook of Social Psychology,* McGraw-Hill, 1998, 143-179.
42. Smith, A. & Yetim, F. Global Human-Computer Systems: Cultural Determinants of Usability [Special Issue], *Interacting with Computers*, *16*, 1 (2004).
43. Sun, H. Exploring Cultural Usability, *Proc. IPCC2002*, IEEE Press (2008), 219-330.
44. Tractinsky, N. Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues, *Proc. CHI'97*, ACM Press (1997), 115-122.
45. van Welie, M., van der Veer, G. C., & Eliëns, A. Breaking Down Usability, *Proc. Interact'99*, IOS Press (1999), 613-620.