



**Copenhagen
Business School**
HANDELSHØJSKOLEN

A framework for thinking about the maturity of cultural usability

Torkil Clemmensen

Department of Informatics
Howitzvej 60
DK - 2000 Frederiksberg

Working Paper nr. 02-2009

A framework for thinking about the maturity of cultural usability

Torkil Clemmensen

Department of Informatics, Copenhagen Business School, Denmark

Abstract: Interaction design and usability have become important contributors to economic and cultural development in emergent economies in today's global distribution of the use and production of IT, but research and practice that incorporate cultural and non-western perspectives on software and interactive products are still in their infancy. This chapter presents theory of cultural cognitive styles and standard usability, and a framework for thinking about the maturity of cultural usability. The framework has five levels. Level I concerns the localization of the user interface, level II focuses on the localization of the usability evaluation methods, level III emerges with new user groups, level IV concerns historical changes in the concept of usability itself, and level V deals with managing a complexity of user groups. The chapter uses recent empirical results from studies of culture and usability to illustrate the need for the framework for thinking about the maturity of cultural usability.

INTRODUCTION

Interaction design and usability have become important contributors to economic and cultural development in emergent economies in today's global distribution of the use and production of information and communication technology. Despite this mega trend, research and practice that incorporate cultural and non-western perspectives on usability of software and interactive products are still in their infancy.



Figure 1. Cultural aspects of the design of traffic regulation signs. How much is culturally different? The form of the warning sign? Design-style? Gender roles? Haircuts? Dress and shoes? Bags? Placement of the children? Walking styles? Pictures from *Mecánico* (2009). © 2009 elve.net. Used with permission.

The Danish picture in Figure 1 is a good introduction to the topic of cultural design. It is a sign that can be found in front of schools in Denmark to warn car drivers to lower their speed so they don't run over children who are crossing the road on their way to school. I found the crossing sign on an international website that displays signs like this from all over the world. The owner of this web site is very happy about this Danish sign. In his opinion, this sign is craftsmanship without any compromise. He tells us to look at the legs, observe the ankles, the waist, the neck, etc. He asks us to notice all the very nice little details in this sign (see, *Mecánico*, 2009). So the question is, of course, how much in this sign is culturally specific. You might have a very easy answer to that, but when the question is asked in Denmark, people tend to believe that everything on the sign is or should be the same all over the world. But if you look at some of the other signs in Figure 1, you see some of the differences. For example, you could expect that the shape of such a sign would be universal because car traffic signs are relatively new things in the world, but it is not. You could expect that people everywhere would have feet, but in some places like in China you can see that people apparently do not have feet to walk on (sic). You could expect that everywhere children would walk to school, but in the Netherlands this poor little girl seems to be chased to school by her big brother, or as the website owner (*Mecánico*, 2009) suggests: "Dutch children are in a hurry to absorb more knowledge. Or... are they leaving school?" or as one of the anonymous reviewers of this chapter suggested: "Two children of different ages running to class"? Also, in the Indian sign the child is really in a hurry to get to school. You could expect that everywhere the big one would be in front protecting the small one, but that's also

different in different places. The owner of the web site suggests that in the Cambodian sign we may see the boy wearing a soldier's cap, the girl a helmet? Maybe they are afraid of stepping into a minefield? (Mecánico, 2009).

The rest of the chapter introduces a psychological definition of culture that is relevant for interacting with technology, discusses issues related to cultural and mainstream usability, and presents a framework for thinking about the maturity of cultural usability. The framework has five levels. Level I and II will be developed in greater detail than are the other three levels. The first level concerns the localization of the interface. The second level focuses on the localization of the methods that we use for evaluating the interface. The third level relates to when a new cultural usability actor emerges because we have new user groups. The fourth level suggests some historical changes in the concept of usability itself, and the fifth level deals with managing the complexity of user groups. Through this review, the reader will hopefully get a view of what usability and culture can be.

CULTURE AND THE USE OF INTERACTIVE TECHNOLOGY

Why do we need a multi-level framework for cultural usability? The road sign shown in Figure 1 surely demonstrates that differences in the immediate perception of signals, signs and symbols are one kind of cultural difference which is evident in design. Another kind of cultural difference in design is what you might refer to as a deeper level. Why is the work of the American psychologist, Nisbett (Nisbett, Peng, Choi, & Norenzayan, 2001), who studies cultural psychology, important for cultural usability? Other researchers more directly related to human computer interaction and culture (e.g., Del Galdo & Nielsen, 1996; Hollan, Hutchins, & Kirsh, 2000; Levanthal et al., 1994; Marcus, 2006) or coming from the tradition of human computer interaction and ethnography (e.g., Anderson, 1994; Button, 2000) could have been mentioned. However, Nisbett's work on cultural differences in cognitive style is particularly relevant to cultural usability, because although human computer interaction takes place in a context, a large part of it is mental, and human computer interaction has a strong historical root in psychology (Clemmensen, 2006). Nisbett, on a solid grounding of empirical psychology research, attempts to build a bridge between cultural psychology and mainstream psychology. His work on cultural cognition could possibly be used to do a similar job of bridging between culture and usability. For example, Nisbett showed 16 of his students in America (and then asked his colleague in Japan to show the same to his Japanese students) short 20 second videos with fish swimming around in a little aquarium. Previous to the experiment it was defined that the fish in the front would be the salient objects and the whole thing would be considered the field. This was an experiment about attention given to field, which is one of the most researched areas in cultural psychology. What they did was, after the students had looked at these 20 second videos, they were then given 90 pictures – 45 pictures of which showed the fish that they had already seen on the video and 45 pictures showing them new fish. Of these pictures that they had already seen, some of the fish had a new background, some no background, and some had the same background as before.

The students were then asked to recognize the pictures. The American students tended not to care about the background, and they were equally able to recognize the fish, whatever the background (or no background) that the fish had. In contrast, the Japanese students tended to think that the fish with a new background was a new fish, a new picture. So Nisbett here provided evidence of how people with

one kind of cultural background pay more attention to field than do people coming from another background. Nisbett's research is based on a distinction between Aristotelian analytical (Western) and Confucian holistic (Eastern) intellectual tradition, which is a controversial distinction. Using analytic thinking includes focusing on the foreground objects, making clear choices between contradictory positions, and more, while using holistic thinking mode means being relational and taking the whole field into account and accepting that there may be some truth in both of two contradictory statements, see e.g. (Nisbett et al., 2001). However, across this cultural distinction Nisbett's experiments on inference patterns help demonstrate that, to understand people's ways of thinking in different cultures, ethnographic observation should be complemented with experimental work (Sperber & Hirschfeld, 2004). Unfortunately, this author has not been able to find any experimental studies of how people coming from societies with other or more mixed intellectual traditions, such as India, react to field or background.

Nisbett et al. found similar results with other experiments on differences between the salient object and the field. They asked students to explain how animals behaved like the fish discussed previously. If you asked the students why these fish moved around like they do, the Chinese students would tend to describe the environment, the plants and the water, etc., while the American students would say it was because it was either a very speedy or a very slow fish. They also asked students to explain about physical events and found similar inference patterns; for example, having a ball in water – why would it go up and down? Chinese students tended to describe the salt in the water making the ball go up. The American students thought it was due to the ball itself being a very light or a very heavy ball.

Further experimental studies carried out by Nisbett et al. concerned how people explain social events. How can we explain unfortunate events, such as when a student enters a university and shoots a number of his fellow students? Nisbett and colleagues looked at how Chinese students and American students explained these stories when these kinds of stories appeared in newspapers. What they found was that the Chinese students tended to argue that it was due to the bad family or bad friends that this student had become a murderer; however, the American students would tend to argue that this student was a murderer because he had a bad personality or something was wrong inside his person.

Thus, in many areas of life we seem to have some really deep and basic differences in our thinking styles, depending on the society we have been raised in. But, of course, this is only true to some degree, because many of us have more than one cultural background. Some people may have two or more cultural backgrounds. Hong and Mallorie (2004), in their dynamic constructivist approach to culture, suggest a situation X cultural model. When two cultural groups like Y and Z, which on average are very different, are compared, then in situation A the difference between Y and Z can be much bigger, while in situation B one can see that there is clearly no difference; in situation C again the difference can be smaller, and in situation D the difference may even be reversed. For example, a Chinese student or an Indian student who has been studying in America, when he encounters some kind of test in the US, might react either as an Indian student or American student, depending on how he is primed, and depending on with whom he is interacting, what he is looking at, which language the

test is in, etc. To some degree, he or she may even consciously activate the cultural background that is preferred in different situations.

In psychology, culture has sometimes been treated synonymously with race or ethnicity as general explanatory factors for intergroup variations in psychological phenomena (Betancourt & López, 1993). This is however too simplified, as race is a biological concept that does not say much about how you live your life. In a similar way, if you happen to be a member of an ethnic minority (e.g., being an immigrant from Turkey in Denmark), this does not necessarily mean that you live your life differently from that of the majority culture (like most Danes). And there are a host of other socio-economic factors that may influence measurable differences in peoples' behavior when they interact with their environment.

Furthermore, what are the apparent and measurable differences in peoples' cognitive style when they interact with their environment? For example, the way they interact with computers in a work situation may depend on who is doing the measuring. What is seen as cultural differences depends on who is looking; thus, culture is in this sense relative. It is therefore important that we as perceivers of cultural differences make it clear why we are looking for these differences and what we intend to do with our observations (Kahn, 1989). For example, coming back to the example in Figure 1 with the Danish road sign, a cultural difference can be identified as being entertaining (e.g., an audience of design students or readers), in which case it is merely a funny anomaly, or it can be seen as the initiation of a process leading to a better design of road signs, in which case the cultural difference is a design opportunity.

Summing up, we can say that a cultural diversity in people's use of interactive designs may be found both in the immediate perception and on a deeper cognitive level. The emergence of cultural differences depends on the technology-use situation. Having a theory, or a framework that defines what cultural differences one is looking for to explain technology use, and knowing the purpose of the identified cultural differences, is therefore important. Without attempting to make an exhaustive list, several different contexts will have to be taken into account when assessing the significance of theoretically and empirically identified cultural differences in the use of technology.

MAINSTREAM USABILITY AND CULTURE

Current standard definitions of usability provide at least four ways of looking at usability (Bevan, 2001, 2006). First, there is usability as quality-in-use or the use of the product (effectiveness, efficiency and satisfaction in a particular context of use). Second, there is the product quality or quality of the user interface and interaction (focus on the product itself, instead of focus on use context). Third, there is the quality of the process used to develop the product. Finally, there is the capability of an organization to apply user centered design. Each of these four standard views on usability may be susceptible to cultural influences, but hitherto little research has been done, and there are a number of problematic issues related to the research that has been done.

Usability as quality in use

Usability as quality in use is often explained with reference to an international standard that says it consists of "... the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in specified contexts of use.." (ISO/IEC FDIS 9126-1, cited in (Bevan, 2001)). This definition implicitly includes cultural differences in the term 'context of use.'

Quality of use has been studied from a cultural point of view. Barber and Badre (1998) argued explicitly that "usability must be re-defined in terms of a cultural context." In 168 native-language websites across 18 countries and nine genres of websites, they identified 16 different web page design elements that proved to be highly prevalent within a country and less prevalent in other countries. The results indicated region specific preferences, for example, a preference for certain navigation elements in the Middle Eastern countries, and also country specific elements such as the multi-coloredness of the Brazilian web sites. In this approach the design elements were seen as a set of cultural markers that could indicate the context of use, which, in turn, was defined as the country/region combined with the genre of the website. The authors' ambition was to develop search engines that parsed the HTML of websites, aggregated cultural markers and assessed their relative frequency according to country of origin, language, and website genre, and in this way, in terms of cultural markers, specified the context of use.

Honold (1999, 2000) illustrated that differences in how people with different cultural backgrounds relate to technologies affect how technologies are actually used. Honold (1999) used focus groups and usability tests in English with local moderators to show 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 rejected the strategy of asking questions to solve problems related to the 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. In this study, the users' learning style provided a context of use. In another study, Honold (2000) observed and interviewed the use of a washing machine in 35 Indian households in Mumbai and Delhi for a three week period. The results showed, among other things, that household users had mental models of top-loader washing machines that were different from that of the front-loader machines in the study, that the machine was not ready for interruptions in electricity common in Indian infrastructure, and that the machine had to be movable for cleaning. In total, she identified eight generic factors for requirement definition in different cultures. On the basis of these factors, she extended a classical activity theory framework (a framework that focus on how tools and artifacts mediate human activity and consciousness, see e.g., Bedny and Meister (1997)), with three more elements: mental models, environment and infrastructure. Honold argued that for each product and each culture this framework should be applied to specify the context of use.

De Angeli et al. (2004) carried out a contextual inquiry study of the adaptation and use of ATMs in Mumbai that involved observations and interviews with 43 early ATM adopters, non ATM users and first time ATM users. The cultural theory used was Hofstede's cultural dimensions, which says that

any country can be located on the four dimensions of masculinity, uncertainty avoidance, power distance, collectivism, and for some countries time-orientation may also be important (Hofstede, 2001). Five specific cultural traits that could affect the adoption of ATMs in urban India were found, of which three were similar to Hofstede's national cultural dimensions (Hofstede, 2001), but two were not. First, communication boundaries that included the use of English in domains such as banking and technology and the problem of poor translation from English to Hindi, and second, the generally low exposure to interactive technology among Indian users at the time of the study. On this basis, the authors suggested that Hofstede's dimensions were of too high a level to inform design of interactive technology, and should be supplemented with empirical research within each emerging market to reveal the context of use for each product in that market. Summing up, together these studies indicate that cultural issues are important for usability as quality in use and, in particular, important for the specification of the context of use.

Usability as Interface quality

Usability seen as product or interface quality can be distinguished from usability as quality in use by focussing on the product itself instead of focussing on the use context. Within the view of usability as interface quality, we find interface programming, product design, and more. Interface quality has been studied from a cultural viewpoint, and an important issue when studying culture and user interface quality has been the language used in the interface.

Dong and Salvendy (1999) carried out usability experiments with horizontal and vertical menus for the Chinese population. They used local moderators, local language and local material systematically. There were 80 Chinese students distributed into four groups: 1) Chinese language interface, Chinese language task instructions, 2) Chinese language interface, English language task instructions, 3) English language interface, Chinese language task instructions, 4) English language interface, English language task instructions, and compared with a group of 20 US students with English language interface and English language task instructions. The results showed that the menu design should be either vertical or horizontal, according to the language of the interface and task used in the evaluation, and not according to the traditional language of the user. This study shows that the language used in the interface influences the quality of the interface.

Native English speakers, however, often have the advantage of superior mastery of the language of the interactive software product and the communication related to such a product. Piamonte (2001) did a questionnaire study of small graphical symbols in video telephony with 48 students in Sweden and 40 students in US. Both groups were given the questionnaires in the English language. The results showed, first, that without sufficient learning opportunities and learning aids, symbols may be harder to understand by almost anyone than previously thought. Second, the results showed that compared to American users, Swedish users tended to give lower certainty ratings, especially during the spontaneous identification of symbols. A possible explanation of this may be found in the fact that American users could answer in their native language, Swedish users had to use English; this difference may possibly have led to lower certainty in recognizing the graphical symbols in the interactive product among the users asked in their non-native language (the Swedish users).

Furthermore, when English is the official language of the society but the ethnicity of the user population is non-English, the cultural specificity of the interface may not be as straightforward as it appears. Chan and Lee (2005) in a study with 72 Hong Kong Chinese users performed usability tests and questionnaires. Hong Kong has English as an official language. The claim by Chan and Lee was that the Hong Kong Chinese population's interaction with technology would be similar to Mainland Chinese users. Chan and Lee (2005) concluded that larger font types, character size, and line spacing, as well as black characters on white background (compared to white on black background), all increased reading speed. Only character size influenced comprehension scores. The assumption behind these results was that the reading purpose was to understand the overall message, in contrast to e.g., detailed proof reading. However, Ji et al.(2004) studied differences in reasoning styles among bilingual Chinese and English speakers, and found that the language of the test had a different effect in terms of cognitive style among different Chinese groups – the use of Chinese as test language was sufficient to bring out a significant preference for relational categorization among Mainland Chinese, but not among Hong Kong Chinese. Hong Kong Chinese users simply did not show similar preferences for relational cognitive style. In the case of the Chan and Lee (2005) study, this could mean that the found effects would be significantly larger in the mainland Chinese population, but we do not have support for this without empirical studies. Findings from studies that use Hong Kong Chinese as representatives of Mainland Chinese users, such as the Chan and Lee (2005) study, should therefore be interpreted with care, and with consideration of possible effects of the language effects on the users' style of reasoning during interaction with technology.

Finally, when English is the official language of society but the user population is obviously multicultural, the cultural specificity of the interface quality is, at best, complicated to explain. Onibere et al. (2001) interviewed 324 users (a variety of 'typical computer users') in a unique study of whether local symbols are more acceptable to users than are standard universal icons, and also whether the current phrases used in menus and various icon descriptions are clearly understood by the up to 15 different local communities with different dialects in Botswana. They found that users preferred localized interfaces, but not necessarily localized icons, and that users did not agree as to which language to use for text-based interfaces. All the interviews were conducted by local interviewers in the English language. The results appeared to be related in complex ways to the use of different languages in different parts of the country, and to the fact that English was the business language of the country. From the point of view of research in culture and usability, this relation between the complexity of multi-language use and product quality of interactive products is very challenging.

Usability as process quality

A view of usability related to the process used to develop the product is expressed in standards concerning a human centered design process (ISO 13407) and usability methods supporting human centered design (ISO 16982) (Bevan, 2001). The standard human centered design process consists of an iteration of activities such as understanding and specifying the context of use and the user/organizational requirements, producing design solutions, and evaluating designs against user/organizational requirements. A key element in the process is to apply usability methods

supporting human centered design such as: think aloud usability testing in which the evaluators have users continually verbalize their thoughts during their activity when using the system, and interviews in which the evaluators ask the users a series of questions about the interface in a more or less structured form.

Culture is not mentioned in standards for usability methods supporting human centered design. For example, interviews as a usability method may be described as having the advantage of getting a quick overview of the users' opinions, and being flexible and allowing probing per users' responses (ISO 16982). Nowhere in usability standards are the advantages of interviews described as being relative to culture. However, Vatrapu and Pérez-Quiñones (2006) in their study of a website for a voluntary non-profit organization used both foreign and local moderators (one Anglo-America and one Indian moderator) and 16 target users in a foreign country (Indian graduate students in US), and found substantial effects of culture on structured interviews. The results showed that participants found more usability problems and gave more suggestions to an interviewer who was a member of the same (Indian) culture than they did to the foreign (Anglo-American) interviewer. The results of the study empirically established that cultural differences among interviewer and interviewees significantly affect the efficacy of structured interviews during international user testing.

Think aloud usability testing is another example illustrating that culture is not discussed in standard views of usability methods, but has been shown by recent research to be prone to cultural influence. The standard view has it that concurrent verbalization, when the user speaks her thoughts while working with the system, is better than retrospective verbalization, when the user speaks after task completion; this is because looking back on what one has done usually allows for selective memories and re-interpretations of what was done and why it was done. Furthermore, an advantage of think aloud test is that it supposedly gives objective, universal access to users' mental processes. However, Yeo (1998) in a think aloud usability testing experiment with Malayan national language spreadsheet users in a Malaysian university found that users with rank higher than that of the evaluator gave more critical comments to the software than users with a lower rank than that of the evaluator. Second, users preferred to think aloud in English mixed with the Malaysian national language. These findings suggest the need for evaluators to have an equal or lower rank as the users, and to be competent in the users' languages. Clemmensen et al. (2009) used Nisbett's work (as discussed in the previous section) to discuss the possible impact of cultural cognition on central elements of think aloud usability testing. Their conclusion, among other things, was that there are different effects of thinking aloud on task performance by people with an Eastern versus a Western cultural background; thus, for Chinese users, retrospective thinking aloud may show more valid results, and cultural differences in nonverbal behavior may also affect the evaluator's reading of the user's mind. The results from the above studies establish, both empirically and theoretically, that there are reasons to believe that think aloud usability testing is influenced by culture.

Not only are there indications that the methods applied in user centered design are prone to cultural influences, but the overall process may also need to be adapted to account for cultural issues. Bourges-Waldegg and Scrivener (1998; 2000) developed their 'meaning in mediated action' framework aimed

at determining contexts shared by culturally diverse user groups, and used these contexts as basis for design representations. The central idea in their framework was to focus on the user's understanding of intended meaning with the designed representation in the given context. They tested the framework with 11 users of an international website, and found supportive indications that this approach can help to find culturally determined usability problems, which improves the international website's usability. Smith et al. (2004) presented a process model for developing usable cross cultural websites that explicitly related theories of culture to practical design activities, and pointed out key concerns such as how to contrast websites with the needs of local users. By doing an analysis of banking websites found in China and India, they exemplified their abstraction of the design process focusing on cultural issues in development.

Usability as the capability of an organization

The view of usability as the capability of an organization to apply user centered design has led to attempts to put forward models for usability maturity. Here, team culture is considered to be the driver that made the development team move from one usability maturity level to the next (for a comprehensive review of usability maturity models for different purposes, (see, Jokela, Siponen, Hirasawa, & Earthy, 2006). Early on, Earthy (1998) developed a scale that measured increasing levels of maturity of human-centered processes. There were five levels, A-E. At the first level of usability maturity, level A, the organization barely recognizes that usability is a problem to be dealt with. On level B the organization has begun to consider the problem of quality in use and has a user focus, and on level C the organization implements user involvement and has usability professionals. On level D, usability is an integrated part of the relevant projects in the organization; finally, on level E usability is institutionalized so that the organization is human centered all the way from top management and down. This usability maturity scale was to be used by usability professionals when they, as consultants, had discussions with client organizations.

The concept of team culture is an important part of the scale; Earthy explains how the most important aspect of an organization moving from one usability maturity level to the next is the culture within the development teams. Thus, moving from A (users are seen as hostile entities) to B (users are seen as people) means going from good practice to a systematical approach; moving from B to C (users are seen as people with different needs than designers) means going from a systematical engineering approach to a user centered engineering approach; moving from C to D (users are seen as formal, theorized entities) means nearly always applying a user centered approach; and moving from D to E (users are perceived as clients/customers/organizations) means going from considering only user centered engineering to seeing this in relation to the clients' business goals. Each of these moves requires a major change in the culture within the development team and its organization (Earthy, 1998).

A related case of making explicit that culture is important for usability is the discussion of the influence of culture on the acceptance of international standards for usability. Bevan (2001, 2006) notes that in the USA there is disagreement about whether to go for national or international standards for usability, while in Europe, the idea of adopting international standards seems more aligned with a

national culture that is influenced by the idea of being strongly associated with many other countries, i.e., with the wish to create common European union standards. Bevan does not explain how culture makes international standards more acceptable in Europe than they are in the USA.

Summing up, we can say that cultural issues are neglected in current standard views of usability. In contrast, recent research shows that: 1) a variety of use contexts may be important for usability viewed as quality-in-use, 2) the language of the interface and use instructions may influence usability viewed as product quality, 3) both usability methods and the overall human centered design process must specifically address cultural issues in order to be valid, and 4) team, organizational and societal culture in complex ways influence an organization's capability to deal with standard usability in a mature way. The problem is how to think about all these issues in a simple and coherent way that allows us to perceive the nature of usability capability of professionals, organizations and other relevant actors to carry out usability from a cultural perspective.

A FRAMEWORK FOR THINKING ABOUT THE MATURITY OF CULTURAL USABILITY

We know now that cultural diversity in people's use of interactive designs may be found both in the immediate perception and on a deeper cognitive level, and that the emergence of such differences depends on the use situation. Without attempting to make an exhaustive list, several different contexts may be important for assessing the significance of theoretical cultural differences in the use of technology. Having a framework (that defines what cultural differences one will be looking for to explain differences in technology use, and knowing to what purpose the identified cultural differences will be put into action) is important. The development of this framework requires taking into consideration the discussion from the previous section, but also broadening it to include the discussion of the possibility of an emerging global IT culture, as well as managing an ever increasing number of diverse user groups. Since the international standard approach to usability does not explicitly deal with the regional, cultural and historical variations in how usability is looked upon in a given society, a framework for thinking about cultural usability should attempt to do that.

The following section presents a framework for thinking about the maturity of cultural usability consisting of five levels. Level I and II will be described in greater detail than the three other levels. The first level concerns the localization of the interface; the second focuses on the localization of the methods that we used for evaluating the interface; the third occurs when a new cultural usability actor emerges because of the introduction of new user groups; the fourth relates to some historical changes in the concept of usability itself; the final level deals with managing the complexity of user groups. This should provide the reader with a clearer view of what cultural usability is. The levels are summarized in

Table 1.

Table 1. The five levels in the cultural usability maturity model

| | |
|-----------|-------------------------------------------------------------------------|
| Level I | National culture as a variable in usability research and development |
| Level II | Getting closer to indigenous usability by localizing evaluation methods |
| Level III | Taking into account global usability culture |

| | |
|----------|-----------------------------------------------------------------|
| Level IV | Usability as an organizing vision for IT development in society |
| Level V | Usability for dynamically changing audiences of users |

A disclaimer is necessary here. This chapter will not be very explicit about the nature of the relations between the levels in the framework. The use of the term ‘Level’ to describe the different pieces of the framework implies a sort of hierarchy or structure. Despite the fact that it may be difficult to see the reasons for this (e.g., the reason to place historical changes in the concept of usability above new user groups) and that less value-laden words such as ‘dimension’ or ‘area’ would perhaps be better descriptors, the term ‘Level’ is used because it help us to keep in mind that there are complex dependencies and overlaps between the different areas of cultural usability, and they are nested within each other. However, being more explicit about the nature of the relation between the levels may be part of a future research agenda.

Level I- National culture as a variable in usability research and development

First of all, cultural usability at level I comprises the traditional way of looking at usability. If you look at culture at that level, usually you look at usability plus some cultural variables in terms of user interface design concerns of particular relevance for internationalization and localization. This is evident from the usability text books that we use to teach our students. Somewhere in the text book there will be a list of cultural variables, which a designer will have to look at to see whether they are relevant. For an example, see Schneiderman and Plaisant (2004, pp30-31, the section on Cultural and international diversity) for a “...long, and yet incomplete... [list of user interface design concerns for internationalization],” which begins with issues such as characters, left-to-right versus right-to-left, date and time formats, etc. The textbook will probably say that considering this list of cultural issues/design concerns/variables will be very helpful if you want to internationalize your design or to localize a design to some specific country. Usability is actually defined at this level, as the standard recommendation is to use existing evaluation methods for user centered designs to test localized designs. Thus, on level I, if a designer is going to develop a user interface for let’s say the Danish audience, he or she will probably consider some variables about how the Danish culture is a feminine culture and individualist culture, and then he or she will probably use some established user centered design approaches. The designer then takes Danish users and tests the interface with these users. Level I in the framework thus encompasses standard definitions of usability as quality in use and as product quality combined with consideration of cultural elements.

Level I cultural usability is closely related to standard ways of doing usability. There are a lot of research results in international research journals that discuss what usability is and how to measure it. There are many studies of different usability methods that indicate which methods are better and which are worse. There are also studies of how many users you need to have in order to identify 80% of the usability problems in a certain software prototype. There are studies of evaluator effects, such as the situation of having different evaluators of a certain user interface indicating the different number and nature of usability problems. There are also studies of usability problem discovery and problem classification that ask what a usability problem actually is, and how such problems should be classified. There are yet other studies on how to communicate these usability problems to the people

who make the changes in the software, i.e., how best to communicate usability findings to designers: give them a report, a PowerPoint presentation, a video, or something else? The aim is to avoid the situation in which designers take a report, throw it under the table and try to forget about it, and then only take it up again when some usability expert reappears in the office. It is important to find ways to keep the image of the user in the mind of the designer, and to communicate the relevant usability problems to the designer; these problems are only multiplied when also considering cultural context and language issues.

Considering the usability problem severity detection rate for different country groups, it can be shown that the evaluator only need five users with a particular set of tests and one target user group in each country to catch 80% of the usability problems on practically any website, though this figure has been disputed by more recent research (Spool & Schroeder, 2001). In real life, however, there are many problems related to detecting usability problems across many countries. Let me offer a personal example. In one of my classes I repeated the study by Law and Hvanberg (2004) of multi-country usability testing. The first problem encountered was language of the interactive product – we wanted to test the Microsoft website for online clipart – and initially I picked three countries: Turkey, Greece and the Middle East regional website for two reasons: these countries and regions are important for Danish trade and tourism, and many students in the class had either traveled to these regions or had family background from there. However, I clearly lacked the cultural skills to evaluate these sites. I could not read the scripts and I could not read the language, which, of course, was a major problem, but perhaps an even bigger problem was that I did not understand the context of use; I did not have qualified explanations of why the images on these websites were different from the images on the comparable Danish website. I would therefore not be able to rate the severity of the problems found in the tests of the sites. Instead, I chose to compare the Danish, Swedish, Norwegian and American websites, which all more or less share the same language, that is, most Danish people can understand and read most of the words and understand the language, and also the US site since most Danish people can also read English. The four sites looked quite similar. I then asked my students to evaluate these sites to identify usability problems. The students worked in groups of two to three students. Within each group one student acted as the evaluator, and the other(s) students acted as test users. Each student group evaluated only one site. For each site there were at least two student groups. The expectation was that we would find the same problems in all sites because there should be the same problems with more or less similar interfaces and same users.

Unexpectedly, in this little exercise with the students, there was an issue with the severity of usability problems. My students tended to rate the problems they found in the Norwegian, Swedish and US sites as severe problems, while the (same) problems found in the Danish website were rated as less severe. Why did they do that? Clearly, they were sort of foreign users because they were not really Norwegian or Swedish or American users. They were Danish students who just looked at some websites in which they were supposed to understand the language and the context of use to a larger degree than they actually did. This small student exercise suggests that foreign users might actually overrate the severity of usability problems that they find in the application. We can call this a “Tourist-effect” in website evaluation, since people apparently behave like tourists and magnify any small

difference that they find between national cultures. The conclusion is that we cannot simply use existing usability evaluation methods in new cultural contexts without considering the cultural bias in the method itself.

Level II - Getting closer to indigenous usability by localizing evaluation methods

At level II the aim is to answer the question of whether usability definitions and usability evaluation methods developed in the US and Europe can actually be used in the rest of the world. This is similar to a view of usability as a process, but emphasizes that the process also needs to be localized.

For example, when considering the think aloud usability test discussed previously, if the person responsible for the usability evaluation works at level II, he or she will ask if the different sub-project usability evaluators do need to have the same cultural background as the test users in order to completely understand how these test users actually will react to and respond to the test instructions and to the test methodology.

At level II, there is a need for a concept of culture other than national culture, a concept of culture that is more focused on individuals' everyday situations and experiences. Level II is about getting closer to indigenous usability by localizing the methods that we use to evaluate usability. If we really want to investigate what people perceive as being good usability, we need our processes and methods to be responsive to their ways of doing things, that is, to peoples' everyday situations, cognitions and experiences. This conception of culture will be helpful when observing people living their daily lives and doing their work. Usability evaluation in that sense will then be localized usability evaluation.

However, apart from the work discussed in the previous section, for example, those by Honold (1999, 2000) and Vatrapu & Pérez-Quñones (2006), there is little research in localizing human centered design and usability methods. For example, how do users apply religious rituals as part of the user software, and how will we evaluate that? We need to have more basic research on target user groups and what good usability is in their life, and to build this knowledge into usability evaluation.

This consideration of characteristics of target users has great practical value for usability processes and methods in offshore software development where there is often a need for using surrogate users, e.g., Indian test users for software designed for the US culture, but also for cross cultural and/or international usability evaluation in general. One example of this kind of consideration from my own research is our field studies on the evaluator – user relation. We (Clemmensen et al., 2007) went to the Chinese website called Baidu (the Chinese Google which has over 50 million Chinese users). They have usability professionals who test their website. We observed such a test and took down notes from that test. One insight we gained was that the evaluator was not really important; he was just a person who gave instructions; thus, the user-task interface relation was the really important element. The user really focused on the task and did only that. Then we went to an Indian company and observed their usability testing. What we saw there was a very caring relationship between the evaluator and the test user; a very polite and very caring relationship. It seemed to be very important how the evaluator

talked to the test users. Finally, we went to a company in Denmark and observed their usability testing. There it was not important at all how the evaluator talked to the test user; instead, the evaluator and, to a certain degree, also the test users were very aware of the client, the ones who paid for the test, but who were actually outside in the observation room. From our research, we saw some possible differences in the communication part of the test in real companies (Clemmensen et al., 2007). It remains for future research to see if these differences hold across many instances of user evaluations in these cultures, and to develop a better explanation of these results culturally.

For a review of theoretical influence of cultural differences on the think aloud usability testing see Clemmensen (2009), In the following paragraphs I will introduce some hitherto unpublished results from our studies of the think aloud usability testing method in diverse cultural contexts. We did explorative quasi-experiments in university labs on the relations between evaluators and users with culturally diverse backgrounds such as age, gender, nationality, mother tongue, familiarity, sexual orientation, etc., while they worked with a localized artifact. The artifact was a clipart collection; the method was the think aloud usability test carried out in Denmark, Assam and Beijing. In each place we used local users. In Assam we had the local students of the Indian Institute of Technology Guwahati, local evaluators from Guwahati and also foreign evaluators – a German and a Dane. Similarly, in China and Denmark we used local users and both local and foreign evaluators. In each place we used a localized version of the clipart to make sure that the application was relevant to the target users who were asked to design an invitation card using the clipart. Among other things, we were interested in seeing if the foreign and the local evaluators doing think aloud usability testing found similar problems with the clipart collection. The test condition with a local evaluator was labeled mono-cultural and the test condition with the foreign evaluator labeled cross-cultural. We did a total of 33 tests, 22 were mono-cultural and 11 cross cultural. As indicated earlier, in a think aloud test, when you ask a user to think aloud while using a piece of software he or she will say incoherent and fragmented utterances like ‘...Oh now I go to... ah, I try to click on this.. Oh, what is this... oh this is a... aha, this must surely be something I had to look at.’ This example of how the user thinks aloud is supposed to be the same all over the world, because task directed information processing is a basic psychological process, that is, every person in the world should be quite good at verbalizing, and, in theory, the process should be similar all over the world. In our experiments there actually were no differences: not only were the local and foreign evaluators similar, but users situated in different geographical places were equally good at thinking aloud in the basic theoretical sense illustrated in the example of thinking aloud above.

However, a large part of people’s verbalizing during a usability test it is not thinking aloud in the strictest theoretical sense, but more explanations or assessments like ‘...I really do not like this interface...This interface is an example of a very poor design...You should really remake this interface if you are the designer...’. This kind of verbalizing is not valid thinking aloud because when giving opinions draws on long term memory the verbalization will be prone to effects of selective memory and post-event-rationalization. With regard to this kind of verbalizing, however, we found a big difference between users with local evaluators who gave a lot of these opinions, criticisms, reflections, explanations, etc., compared to users with foreign evaluators who did none of this. The

difference could be related to the fact that the local evaluator actually helped out the user a lot, that is, if the user was stuck in the task and could not figure out what to do next, the local evaluator tended to help the user, while the foreign evaluator did not do that, but remained passive and did not offer any help. Furthermore, the evaluators' use of affirmatives was different: the local evaluators tended to encourage the local users by saying 'yes, go on, it is very good, go on keep working, keep thinking aloud', etc., while the foreign evaluators would say nothing or very little during the test.

Do these differences between mono- and cross cultural think aloud test conditions have any consequences for design of IT? In our quasi-experiments we found 35 different kinds of usability problems. Of these, nine were found only in the local evaluator (mono cultural) test condition. Apparently, there is something about the local evaluator-local user situation which helps them find some unique problems with a localized software.

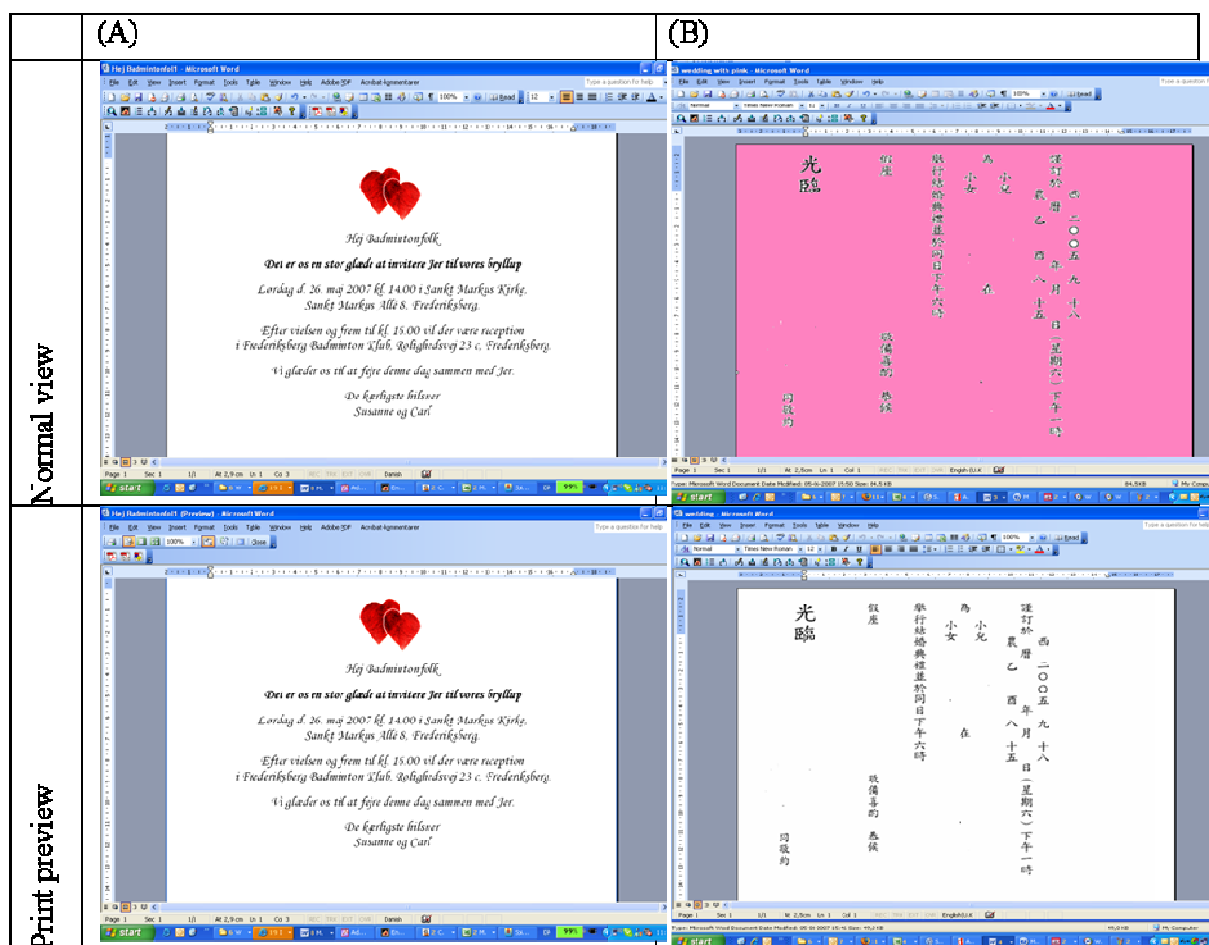


Figure 2. A usability problem found in a quasi-experimental test condition in Beijing. The problem was found only in the local evaluator-local user test condition, and not in conditions where the evaluator was foreign. Disadvantage of Chinese style users due to lack of print preview function with background color; see column B normal view (upper row) and print pre-view (Lower row). For comparison, a Danish style invitation is shown in column A.

The following is an example of a problem which was only found in a local evaluator-local user test condition. The users were asked to make a wedding invitation. In Figure 2 on the left you see the Danish wedding invitation and in the front some little clip arts of some hearts – this is common in Denmark. On the right side of Figure 2 one can see the Chinese wedding invitation. Chinese users tend to also consider the background style, i.e., the paper used for the wedding invitation is important, so in our tests they focused on getting the right (pink) color of the paper. The situation is that in the word processing software that the clipart collection was embedded in, Microsoft Word 2003, there is a nice print-preview function. If you use that function, you get a perfect preview of your printed wedding invitation for the Danish users, because you see the same picture in preview as that which you have designed in the word processor. But as a ‘background style user’, e.g., in our case Chinese users, you have a problem because while you can see the background color in your word processor, the preview does not show the color of the background. This is also true in the standard configuration of Microsoft Office 2007. This is an example of a cultural bias in a very well known piece of office software. It was found only in a test situation with a local test user. I do not know why, but it did not appear in the situation of the foreign moderator. It may be because the users felt that it would be too difficult to explain to the foreign moderator why seeing the background color was a very important thing, or it may be that the foreign evaluator did not accept the missing background-preview as a problem.

Level III – Taking into account global usability culture

Due to technological developments, the concepts of usability and culture may merge into a new level of cultural usability, level III. For example, the Nokia user segmentation model may in principle be independent of geographical information, i.e., it has to do only with user segments such as ‘the elderly user’, or the ‘very technologically advanced user’, etc., and these segments are the same all over the world. It seems that in our new global world new user groups appear, and that we are no longer talking about the national culture or the local culture, but about the technology driven cultures, the IT cultures. Despite this image of a global IT culture, in real life, in practice, Nokia had to do a certain kind of studies of the local user groups too, and add this kind of national culture level information to their global segmentation of users (personal communication with Nokia user experience manager, 2008). However, new technologies like mobile phones create new user groups and new IT cultures across national borders around the world, and we need to strike the right balance between talking about local traditional culture in the home and in the workplace, and talking about the new IT culture, the development of which may be driven by new consumer products.

Level IV – Usability as an organizing vision for IT development in society

The fourth (IV) level of usability is related to the spreading of the concept of usability in society. At a certain point in time after the initial introduction in society, usability becomes ‘folk science’, and everybody uses the word usability now and then. In Denmark, the usability concept appeared in 1984. Since then it very slowly spread in society and today it is such a common concept that everybody knows usability - it can be used in all kind of contexts - people use it in their everyday language and also outside information technology, to talk about the usability of a newspaper, of hairdryers, etc. But it is something that took a number of years in Denmark. In comparison, usability in, for example, China or India, has spread very rapidly, and within the last five to ten years many people know this

concept. An important type of studies of cultural usability needed to develop our understanding of Level IV would be looking at newspapers and IT magazines and analyze how the talk about usability changed historically from, for example, the year 2000 until today. In a given society, how has the way people talk about usability changed? What would happen if you really decided on a national basis, for example, in India or China, that usability should become an organizing vision of IT development? This would require that government actively supported usability in education, administration, and business.

Level V – Usability for dynamically changing audiences of users

Level V is when usability is related to dynamically changing audiences of users. While Level III is about different user groups in different places in the world and the possible new world-wide IT user groups, and Level IV is about the spreading of the concept of usability, Level V is about the situation in which the local users have many, but different cultural backgrounds. At level V, usability is not considered to be a stable trait of technology or user population, but rather a dynamic mode of interaction. This is when usability is seen as something that, although most often relatively stable, can change, and something which a large diversity of users, rather than merely passively experience, actively construe. This happens in many places in the world where for a long time there has been a great cultural diversity in the population, such as in big cities in India. This, however, is a situation that is only emerging in many European countries.

Denmark can therefore be used as an example of the emergence of this level of cultural usability. First, Denmark is increasingly becoming multi-cultural. Today, 10% of the population is immigrants or children of immigrants, a number which is increasing steadily. However, this has only changed within the past 15 years. The Danish society is currently less homogenous compared to when the concept of usability was introduced in the early 1980s, so design for usability should reflect the new diversity in user groups. Second, since an important European treaty was decided in 1992, Denmark has begun to become a part of the European Union not only in business, but also in other parts of life, so that Danish students and workers can go abroad, and foreign workers and students can come to Denmark in increasing numbers. This means that usability of technology will have to be related to common work and study situations in which within a few hours, days or weeks the participants will have to deal with several human languages such as English, German, Danish, Italian, etc. Third, because the use of computers became common in the early 80s in Denmark, most people today live surrounded by computers both at home and at work in all or most of their lifetime. Also, elderly people adopt and use the newest technology, so young people in Denmark are definitely not the only group that uses new technology. Fourth, Denmark is overwhelmingly a Christian society and there is little awareness about how to design for other religious groups. In Denmark, people are not used to the situation which is common in many other places around the world – actually designing for a very diverse audience of users, of which many are able to shift instantly between their several, different cultural backgrounds.

FUTURE MATURITY LEVELS OF CULTURAL USABILITY

Looking to the future, recent research indicates a trend towards identifying indigenous usabilities, e.g., replacing the windows metaphor with a Chinese garden metaphor. If this trend is to be supported, there is a pressing need for identifying research on indigenous psychologies that can be used as a theoretical basis for such indigenous usability. Much of existing cultural psychology compares the US and China, but in this East-West dichotomy, where, for example, would India be located on the scale? Regarding the polite and sensitive relationship that seems to exist in the usability test situation in India, is this preference related to the Indian society and ways of doing things with technology? There is research on Indian psychology and Indian perspectives on cognition (Mishra, 2006), which seems to be indicating that in India, the context of use is of highest importance when defining usability. Future research could try to identify more of this research on indigenous psychology and combine that with usability to create indigenous usabilities in different cultures.

If we accept that micro-concepts of culture, such as cognitive style, influence technology use in different countries/user groups, then new questions arise. Which parts of the different indigenous usabilities overlap with main stream usability views, and which do not? Are styles of interaction with computers value free or value laden? Are some cognitive styles simply better when interacting with technology (e.g., whether it is better for an aircraft pilot to be field independent)?

Learning from the previous discussions of culture and IT, it seems that in the future we have to deal with usability more as a mode of interaction than as an invariant trait of specific user groups, use situations or products. It will become more important to identify not only the variety of styles of interactions with computers that are possible around the world, but also how people shift and choose in different situations between different styles of technology use available to them.

I perceive the future of usability as constantly having to think about multi cultural users, multi lingual users, and global users. You have them in the same user group, and their interactions with technology may change both rapidly and over time. You can have people who in one situation act according to their Arabic background, and the next moment they act according to their Danish background. Many of my students have multicultural backgrounds, and I keep being surprised by their cultural competences and fluent abilities in diverse cultural contexts. How can one design for such a challenge?

CONCLUSION

This chapter has introduced cultural usability and then addressed cultural theory relevant to interactive technology use, with an emphasis on cultural cognitive psychology. A cultural diversity in people's use of interactive designs may be found both in the immediate perception and on a deeper cognitive level. The emergence of cultural differences depends on the usage situation. Having a theory (or a framework defining cultural differences that you will be looking for to explain technology use, and knowing to what purpose you will put the identified cultural differences into action) is therefore important. Without attempting to make an exhaustive list, several different contexts will have to be taken into account when assessing the significance of theoretically and empirically identified cultural differences in the use of technology.

We then analyzed standard usability definitions in relation to research on culture, and found that cultural issues are often neglected in current standard views of usability. Recent research shows that a variety of use contexts may be important for usability viewed as quality-in-use. The language of the interface and usage instructions may influence usability viewed as product quality. Both usability methods and the overall human centered design process must specifically address cultural issues in order to be valid. Team, organizational and societal culture in complex ways influences an organization's capability to deal with standard usability in a mature way. The problem is how to think about all these issues in a simple and coherent way that allows us to get a picture of usability professionals', organizations' and other relevant actors' capabilities to view usability from a cultural perspective.

Finally, this chapter has presented a framework for thinking about the maturity of cultural usability in five levels. Level I and II were described in a more elaborate way than the three other levels. Level I concerns the localization of the interface; Level II focuses on the localization of the methods that we use for evaluating the interface; Level III emerges when a new cultural usability actor emerges because we have new user groups; Level IV concerns historical changes in the concept of usability itself; and level V is about designing for, and managing, a complexity of user groups.

Suggested future research on frameworks for maturity cultural usability includes identifying research on indigenous psychologies that can be used as a theoretical basis for indigenous usability and, in this way, help assess the maturity of cultural usability. Similarly, it would be helpful to assess whether some cultural cognitive styles are simply better when interacting with technology. It would also be useful to explore how people shift and choose, in different situations, between different styles of technology use available to them.

Among the limitations of this work is that cultural cognitive psychology is only one among many relevant perspectives of culture interacting with technology. Other perspectives on culture and IT that could be taken into consideration are contextual design and enquiry, anthropology and ethnography of human computer interaction, and the work in Information Systems research on culture and IT.

Acknowledgements

This study was co-funded by the Danish Council for Independent Research (DCIR) through its support of the Cultural Usability project.

REFERENCES

- Anderson, R. J. (1994). Representations and requirements: The value of ethnography in system design. *Human-Computer Interaction*, 9(3), 151-182.
- Barber, W., & Badre, A. (1998). *Culturability: The Merging of Culture and Usability* Paper presented at the 4th conference on Human Factors and the web, New Jersey, USA.
- Bedny, G., & Meister, D. (1997). *The Russian theory of activity: Current applications to design and learning*: Lawrence Erlbaum Associates.
- Betancourt, H., & López, S. R. (1993). The study of culture, ethnicity, and race in American psychology. *American Psychologist*, 48, 629-629.

- Bevan, N. (2001). International standards for HCI and usability. *International Journal of Human Computer Studies*, 55(4), 533-552.
- Bevan, N. (2006). International Standards for HCI. In *Encyclopedia of Human Computer Interaction*: Idea Group Publishing.
- Bourges-Waldegg, P., & Scrivener, S. (1998). Meaning, the central issue in cross-cultural HCI design. *Interacting with Computers*, 9.
- Bourges-Waldegg, P., & Scrivener, S. A. R. (2000). Applying and testing an approach to design for culturally diverse user groups. *Interacting with Computers*, 13(2), 111.
- Button, G. (2000). The ethnographic tradition and design. *Design Studies*, 21(4), 319-332.
- Chan, A. H. S., & Lee, P. S. K. (2005). Effect of display factors on Chinese reading times, comprehension scores and preferences. *Behaviour and Information Technology*, 24(2), 81-91.
- Clemmensen, T. (2006). Whatever happened to the Psychology of Human-Computer Interaction? a biography of the life of a psychological framework within a HCI journal *Information Technology and People*, 19(2), 121-151
- Clemmensen, T., Hertzum, M., Hornbæk, K., Shi, Q., & Yammiyavar, P. (2009). Cultural cognition in usability evaluation. *Interacting with Computers*, 21(3), 212-220.
- Clemmensen, T., Shi, Q., Kumar, J., Li, H., Sun, X., & Yammiyavar, P. (2007). Cultural Usability Tests -- How Usability Tests Are Not the Same All over the World. In N. Aykin (Ed.), *Usability and Internationalization - Second International Conference on Usability and Internationalization, UI-HCII 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007, Proceedings, Part I* (Vol. 4559, pp. 281-290). Heidelberg: Springer.
- De Angeli, A., Athavankar, U., Joshi, A., Coventry, L., & Johnson, G. I. (2004). Introducing ATMs in India: a contextual inquiry. *Interacting with Computers*, 16(1), 29.
- Del Galdo, E. M., & Nielsen, J. (1996). *International user interfaces*: Wiley New York.
- Dong, J., & Salvendy, G. (1999). Designing menus for the Chinese population: horizontal or vertical? *Behaviour and Information Technology*, 18(6), 467-471.
- Earthy, J. (1998). Usability maturity model: Human centredness scale. *INUSE Project deliverable D*, 5.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*: Sage Publications.
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2), 174-196.
- Hong, Y.-Y., & Mallorie, L. M. (2004). A dynamic constructivist approach to culture: Lessons learned from personality psychology. *Journal of Research in Personality*, 38, 59-67.
- 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). Cultural 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.
- Ji, L. J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, 87, 57-65.
- Jokela, T., Siponen, M., Hirasawa, N., & Earthy, J. (2006). A survey of usability capability maturity models: implications for practice and research. *Behaviour & Information Technology*, 25(3), 263-282.
- Kahn, J. S. (1989). Culture: demise or resurrection? *Critique of Anthropology*, 9(2), 5.
- Law, E. L.-C., & Hvannberg, E. T. (2004). *Analysis of Combinatorial User Effects in International Usability Tests*. Paper presented at the CHI 2004, Vienna, Austria
- Levanthal, L., Teasley, B., Stone, D., Lancaster, A. M., Marcus, A., Nardi, B., et al. (1994). *Designing for diverse users: will just a better interface do?* Paper presented at the CHI 1994.
- Marcus, A. (2006). Culture: Wanted? Alive or Dead. *Journal of Usability studies*, 1(2), 62-63.
- Mecánico, B. (2009). Some common traffic signs of the world - Children on the road in depth. Retrieved August 1, 2009, from <http://www.elve.net/rkid02.htm>

- Mishra, R. C. (2006). Indian Perspectives on Cognition. In U. Kim, K.-S. Yang & K.-K. Hwang (Eds.), *Indigenous and Cultural Psychology- understanding people in context*: Springer.
- Nisbett, R. E., Peng, K. P., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, 108(2), 291-310.
- Onibere, E. A., Morgan, S., Busang, E. M., & Mpoeleng, D. (2001). Human-computer interface design issues for a multi-cultural and multi-lingual English speaking country - Botswana. *Interacting with Computers*, 13(4), 497-512.
- Piamonte, D. P. T., Abeysekera, J. D. A., & Ohlsson, K. (2001). Understanding small graphical symbols: a cross-cultural study. *International Journal of Industrial Ergonomics*, 27(6), 399-404.
- Schneiderman, B., & Plaisant, C. (2004). *Designing the User Interface* (Fourth edition ed.): Pearson Addison Wesley.
- Smith, A., Dunckley, L., French, T., Minocha, S., & Chang, Y. (2004). A process model for developing usable cross-cultural websites. *Interacting with Computers*, 16(1), 63.
- Sperber, D., & Hirschfeld, L. A. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Sciences*, 8(1), 40-46.
- Spool, J., & Schroeder, W. (2001). *Testing web sites: five users is nowhere near enough*. Paper presented at the CHI '01 extended abstracts on Human factors in computing systems.
- Vatrapu, R., & Pérez-Quñiones, M. (2006). Culture and Usability Evaluation: The Effects of Culture in Structured Interviews. *Journal of Usability studies*, 1(August), 156-170.
- Yeo, A. W. (1998). *Cultural Effects in Usability Assessment*. Paper presented at the CHI 98, Doctoral Consortium.