Improving Performance, Perceived Usability, and Aesthetics with Culturally Adaptive User Interfaces

KATHARINA REINECKE and ABRAHAM BERNSTEIN, University of Zurich

When we investigate the usability and aesthetics of user interfaces, we rarely take into account that what users perceive as beautiful and usable strongly depends on their cultural background. In this paper, we argue that it is not feasible to design one interface that appeals to all users of an increasingly global audience. Instead, we propose to design culturally adaptive systems, which automatically generate personalized interfaces that correspond to cultural preferences. In an evaluation of one such system, we demonstrate that a majority of international participants preferred their personalized versions over a nonadapted interface of the same Website. Results show that users were 22% faster using the culturally adapted interface, needed fewer clicks, and made fewer errors, in line with subjective results demonstrating that they found the adapted version significantly easier to use. Our findings show that interfaces that adapt to cultural preferences can immensely increase the user experience.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*User-centered design*

General Terms: Human Factors, Design, Experimentation

Additional Key Words and Phrases: Cultural adaptivity, localization, MOCCA, aesthetics, usability, user experience

ACM Reference Format:
Reinecke, K. and Bernstein, A. 2011. Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces. ACM Trans. Comput.-Hum. Interact. 18, 2, Article 8 (June 2011), 29 pages.
DOI = 10.1145/1970378.1970382 http://doi.acm.org/10.1145/1970378.1970382

1. INTRODUCTION

With the exponentially growing number of Webpages on the Internet, searching for the best information increasingly becomes a search for the best information *presentation*. Many sites simply contradict one's personal understanding of "good design." But even worse, bad design often goes hand in hand with bad usability. Such findings have led to a growing awareness that aesthetic design is a key component of usability [Hassenzahl 2004; Norman 2004], and a decisive aspect for marketplace success [Bloch 1995]. If neglected, many users rightly decide on another, more attractive and usable Website offering similar content [Lindgaard and Dudek 2003]. With this in mind, research has long discussed the magic formula for a "perfect design" and tried to define what is perceived as beautiful and usable.

To some extent, it seems that we can indeed generalize what users consider usable and attractive (e.g., when designing according to the laws of Gestalt Psychology). Some

K. Reinecke is currently affiliated with Harvard School of Engineering and Applied Sciences.
The authors thank the Hasler Foundation for partly funding this research under grant no. 2322.
Author's address: K. Reinecke: email: reinecke@seas.harvard.edu.

DOI 10.1145/1970378.1970382 http://doi.acm.org/10.1145/1970378.1970382

aspects, however, are certainly a matter of personal taste [Norman 2004], or strongly influenced by cultural values [Ito and Nakakoji 1996]. The same conventions apply to user interface preferences, which highly vary across cultures: Asian Websites, for instance, are commonly bright and colorful, with frequent animations that try to attract the user's attention. Their high complexity is often perceived as information overload by Westerners, who prefer more structured content [Marcus and Gould 2001]. Asians, in contrast, have been shown to efficiently filter such dense information [Nisbett 2003].

Users also differ in their design preferences and in their perception of usability at the country level [Barber and Badre 1998; Callahan 2005]. The importance of considering such culturally determined partialities for a certain look & feel of user interfaces has been demonstrated numerous times. For example, interfaces designed for users from a specific country were perceived as much more attractive [Corbitt et al. 2002], and improved the work efficiency of those they were intended for [Badre 2000; Ford and Gelderblom 2003].

Thus, with the Internet being an increasingly global market place, it is unreasonable to design one common Website for everyone, and yet expect to attract an international audience. A more desirable solution would be to design many different versions of a Webpage in order to cater to users' cultural preferences. Unfortunately, in this case, designers would soon experience the problem of scale due to the complex and intangible nature of cultural background.

To bridge this dichotomy between the need for Websites that cater for individual cultural backgrounds, and an inexpensive method to develop them, we have previously proposed an approach called cultural adaptivity [Reinecke and Bernstein 2007]. The idea is that Websites automatically compose personalized interfaces based on the individual cultural backgrounds of its users.

We expect cultural adaptivity to *improve the performance, as well as the user satisfaction* of a personalized web site compared to a nonadapted version of the same site.

We refer to performance as the objectively measurable work efficiency, and to user satisfaction as an umbrella term for the users' subjective perception of the interface, composed of their impressions of both usability, and aesthetics. According to ISO 9241-11 [1997], usability consists of effectiveness (e.g., a user's ability to successfully find information and accomplish tasks on a Webpage), efficiency (he or she can do so in an adequate time frame without frustration), and satisfaction (whether the user enjoys using the Website). Note that rather than being three independently measurable terms, these aspects of usability have been found to interact [Lindgaard and Dudek 2003]. The user's judgement of his or her satisfaction, for instance, is thought to be influenced by the perceived usability (as denoted by the effectiveness and efficiency), but also by affective aspects, such as aesthetics [Lindgaard and Dudek 2003]. In the following, we use attractiveness, appeal, and beauty as synonyms for these affective aspects.

In an evaluation with 41 users of different cultural backgrounds, we demonstrate the benefit of our approach by showing that the majority preferred their culturally personalized interfaces over a nonadapted version. Moreover, our participants were able to work 22% faster with the culturally adapted interface, showing that our approach improves both the performance as well as the user satisfaction.

In the following, we first discuss the related work, before providing some background on our approach to cultural adaptivity in Section 3. Section 3.2 details the implementation of this approach in our culturally adaptive prototype, which was first introduced in [Reinecke and Bernstein 2009]. The experiment conducted with the system is described in the subsequent section, followed by a discussion of the implications of this approach for international users.

2. RELATED WORK

While both industry and research have long realized that adapting Websites to a user's culture is a key to market success, efforts in this direction have been limited to software localization. Localized user interfaces usually modify the most obvious elements to suit a target country and/or region, for instance, by adapting to different languages and regional characteristics [Bourges-Waldegg and Scrivener 1998]. Until today, localization did not consider the adaptation of user interfaces at presentation level, so that in most cases, the interface's complexity, colors, and workflows remain unchanged for all users [Kersten et al. 2002]. Despite this fact, research has acknowledged that providing one user interface per country is not enough [Reinecke and Bernstein 2007], due to the difficulty of pinning down cultural background. Indeed, culture does not keep within "artificial" boundaries of countries [Hofstede 1997]. Contrary to this, research has shown that a national interpretation of culture, where the term is equalized with a specific country, does bundle many user preferences (see, e.g., Baumgartner [2003] for an overview): a fact that we have exemplified with Websites of Asians and Westerners in the introduction of this article. Consequently, many researchers have tried to extract design rules by analyzing and comparing Websites [Yeo 1996; Barber and Badre 1998; Gould et al. 2000; Zhao et al. 2003; Burgmann et al. 2006], often resulting in localization guidelines for specific countries, or regions. The implementation of these guidelines in the design of user interfaces is, however, only beneficial for users who (a) belong to a group targeted by the company, and (b) have not been influenced by other countries (and national cultures). An Indian living in Germany, for instance, might have adopted parts of German culture, and her preferences towards Webpages have possibly changed. Most likely, she will not feel completely comfortable with German user interfaces that often stand out with a plain design, few colors, and minimal complexity. Despite our increasingly global world, our Indian user and other culturally ambiguous people still have to decide between one interface or the other [Reinecke et al. 2010].

Recognizing this problem, researchers have previously proposed to equip computers with some kind of cultural adaptivity (see, e.g., O'Neill-Brown [1997] with her call for culturally adaptive agents). The idea is to develop systems that can automatically recognize and adapt themselves to the user's cultural background. To our knowledge, work in a comparable direction has been limited to research conducted by Kamentz [2006] in the area of e-learning, and by Heimgärtner [2005] in the area of cultural adaptivity in navigation systems. Kamentz [2006] relied on a questionnaire to classify the user into one of a set of predefined cultural groups, and this classification triggered adaptations to an e-learning system. Since the adaptations were mainly aimed at improving the user's learning experience, she focused on the learning style (e.g., an adaptation of the instructions), and symbols. The adaptations did not comprise a full rearrangement of user interface components, as we anticipate is necessary for culturally adaptive software. While our aim is also to overcome some of the disadvantages of localization, such as the previously mentioned disregard of culturally ambiguous users, Kamentz directed her work at only a few countries, and did not include affiliations to more than one country.

Heimgärtner [2005], in comparison, concentrated on the classification of users by their interaction patterns. He introduced a tool, which automatically classifies users based on their navigational patterns while carrying out predetermined tasks. His work has focused on paving the way for culturally adaptive navigation systems [Heimgärtner et al. 2007], and includes different adaptations as they are necessary for culturally adaptive Webpages.

Work towards real-life systems that incorporate cultural adaptivity has possibly been hampered by a fear that intelligent systems might disturb the user [Benyon 1993; Shneiderman 2002]. However, adaptive systems in other disciplines have previously

been demonstrated to improve the performance of users [Greenberg and Witten 1985; Sears and Shneiderman 1994; Höök 1997; Gajos et al. 2008; Findlater and McGrenere 2008; Findlater et al. 2009]. Since cultural background affects the preference for more or less guidance, or more or less complex interfaces [Marcus and Gould 2001; Reinecke and Bernstein 2009], the same could be expected for culturally adaptive interfaces. Additionally, we anticipate that users perceive a personalized version as more aesthetically appealing, as culture highly influences the most obvious visual preferences, such as color [Badre 2000; Kondratova and Goldfarb 2006]. Both points are addressed in the evaluation part of this article, which follows the introduction to our approach in the next section.

3. CULTURAL ADAPTIVITY

Designing systems that automatically adapt themselves to the user's cultural background first requires a deep understanding of the term *culture*, and its effects on user interface preferences. We therefore start with a succinct introduction on culture, before describing how we have chosen to approach cultural adaptivity. A more elaborated discussion can be found in Reinecke et al. [2010].

3.1. Culture and Its Influences on Human-Computer Interaction

While culture has been described numerous times, cultural anthropologists have long agreed that the term cannot be pinned down to a finite definition [Kroeber and Kluckhohn 1952]. Researchers have therefore often outlined aspects that influence culture: For one, this is the national identity someone is born into, which is often equated with the country of origin [Rogers and Tan 2008]. Anthropologists have a more comprehensive understanding; they distinguish between "place," referring to where a person is currently located, and "space," which describes a person's mental affiliation as in the case of cultural ambiguity (simplified from Gupta and Ferguson [1997]). Accordingly, it is not enough to adapt interfaces to the user's current country of origin, as it is common practice in localization. Instead, possible former countries of residence and differing nationalities of the parents may also influence user interface preferences, and therefore, should be taken into account in an approach to cultural adaptivity.

In addition, there are many aspects of culture that have been found to impact interface preferences, such as a user's first (and second) language [Nisbett 2003; Röse 2005], religion [Siala et al. 2004], the education level and the form of education. For example, people who have mostly received teacher-centered instruction are more likely to appreciate detailed instructions later in life [Reinecke 2005], and this seems to be a decisive factor for the self-determined handling of computers.

Other factors are social and political norms, which influence whether people think self-centeredly, or see themselves as part of a group [Schmid-Isler 2000; Nisbett 2003]. As a result to this, people belonging to Eastern cultures find it easier to recognize relationships between several items, whereas Westerners have been found to mainly focus on individual objects [Nisbett 2003].

While the aforementioned aspects describe possible influences on culture, anthropology has avoided defining a finite set of such influences. However, with many disciplines calling for the possibility to compare countries and their national cultures, some researchers have attempted to define so-called cultural dimensions. One of the most comprehensive studies towards such a classification is the one conducted by cultural anthropologist Hofstede [2001]: After analyzing the characteristics of IBM employees in different countries, he distinguishes between the five cultural dimensions Power Distance (PDI), Individualism (IDV), Masculinity (MAS), Uncertainty Avoidance (UAI), and Long Term Orientation (LTO) (see Hofstede [2001] for details). The position a country has obtained within one cultural dimension is denoted by a score, allowing for

comparison between countries. Malaysia, for example, has one of the highest Power Distance scores worldwide (104); a dimension that relates to the perception of hierarchies within society. Thus, in comparison to most other countries, Malaysians are much more tolerant towards an unequal distribution of power, and less powerful members of society accept the lack of democratic rights [Hofstede 2001].

Although Hofstede's dimensions have often been criticized for generalizing the diverse concept of culture, and reducing it to nationality [McSweeney 2002], researchers have applied his dimensions in different disciplines, ranging from interpersonal communication, to human-computer interaction. Much work has been invested in comparing web sites of different countries and finding the relationship between their designs and Hofstede's cultural dimensions (see references in Table I). The results indicate that his dimensions impact user interface designs, and can be used to explain layout differences between various countries. In addition, findings from other disciplines can be interpreted to reveal needs for the interaction with computers. For example, research on teacher-student relationships indicate that students in low Power Distance societies are often expected to learn independently and define their own learning paths, in contrast to high Power Distance societies, where teachers are supposed to define the next steps [Hofstede 1986]. Applied to the context of user interface design, this practice could later lead to a preference for many functionalities, or a reduced choice of options and navigation paths. We have summarized these findings and their implications for user interface design in Table I, which broadly suggests what users might like. In previous work, we evaluated these relations in order to ensure that they form a suitable basis for adaptation rules [Reinecke and Bernstein 2008, 2009].

3.2. An Approach to Cultural Adaptivity

In an approach to cultural adaptivity, the aforementioned influences on culture need to be acquired about each user, stored in a personal user model instance, and mapped onto user interface adaptations.

Thus, as a first step it is necessary to build a user model based on cultural particularities before the adaptation can be accomplished. The idea is that the computer acquires the user's cultural background by taking into account various cultural influences that affect user interface preferences. This information can be saved in a cultural user model ontology, which we previously introduced in Reinecke et al. [2007] and Reinecke et al. [2010]. For the computer, the ontology serves as a knowledge base about each user, with every user model aspect linking to a set of rules, which trigger the adaptation of the user interface. To correct unsuitable adaptations, the user should be able to add more information about his or her cultural background to the user model. Likewise, the application connected to the user model should also be able to learn new, and refine existing adaptation rules. If the user model is updated by either the user (manually) or the computer (automatically through observation of the user's interaction with the system), the adaptation rules are automatically updated too, triggering new adaptations of the user interface.

Figure 1 illustrates this adaptive process in more detail. If the user model is employed for the first time, the user needs to initially provide information in a short questionnaire provided by an application, and/or a user model editor. In our approach, this explicit acquisition of information helps to mitigate the cold-start problem [Mehta and Nejdl 2007], which occurs when little is known about a user who is new to a system. We therefore restricted the questions that users have to answer to only three questions: "Where do you currently live?", "In which other countries have you lived before?", and "How long have you lived in each of these countries?" The answers are added to the user's respective user model instance on the user model server (see top arrow), so that the user's cultural dimensions can be calculated by drawing information from

Table I.
Adaptation rules as derived from related work on the effect of Hofstede's dimensions on user interface design for users with a high or low score

	Low score	High score	Derived from
Power Distance	Different access and navigation possibilities; nonlinear	linear navigation, few links, minimize navigation possibilities	[Voehringer-Kuhnt 2002; Hofstede 1986; Marcus and Gould 2000; Burgmann et al. 2006]
	Data does not have to be structured	Structured data	[Marcus and Gould 2000]
	Many functionalities	Reduced choice of functionalities	[Hofstede 1986]
	Most information at interface level, hierarchy of information less deep	Little information at first level	[Marcus and Gould 2000; Burgmann et al. 2006]
	Friendly error messages suggesting how to proceed	Strict error messages	[Hofstede 1986; Marcus and Gould 2000; 2001]
	Support is only rarely needed	Provide strong support with the help of wizards	[Marcus and Gould 2000]
	Websites often contain images showing the country's leader or the whole nation	Images show people in their daily activities	[Marcus and Gould 2000; Gould et al. 2000]
Individualism	Traditional colors and images	Use color to encode information	[Marcus and Gould 2000]
	High image-to-text ratio	High text-to-image ratio	[Gould et al. 2000]
	High multimodality	Low multimodality	[Hermeking 2005]
	Colorful interface	Monotonously colored interface	[Barber and Badre 1998]
Masculinity	Little saturation, pastel colors	Highly contrasting, bright colors	[Voehringer-Kuhnt 2002; Dormann and Chisalita 2002]
	Allow for exploration and different paths to navigate	Restrict navigation possibilities	[Ackerman 2002]
	Personal presentation of content and friendly communication with the user	Use encouraging words to communicate	[Hofstede 1986; Dormann and Chisalita 2002; Callahan 2005]
Uncertainty Avoidance	Most information at interface level, complex interfaces	Organize information hierarchically	[Marcus 2000; Marcus and Gould 2000; 2001; Zahed et al. 2001; Hodemacher et al. 2005; Cha et al. 2005; Choi et al. 2005; Burgmann et al. 2006]
	Non-linear navigation	Linear navigation paths/show the position of the user	[Marcus 2000; Marcus and Gould 2000; 2001; Baumgartner 2003; Hofstede 1986; Kamentz et al. 2003; Corbitt et al. 2002; Burgmann et al. 2006]
	Code colors, typography and sound to maximize information	Use redundant cues to reduce ambiguity	[Marcus and Gould 2000; 2001]
Long Term Or.	Reduced information density	Most information at interface level	[Marcus and Gould 2000; Marcus and Baumgartner 2004]
	Content highly structured into small units	Content can be arranged around a focal area	[Marcus and Gould 2000]

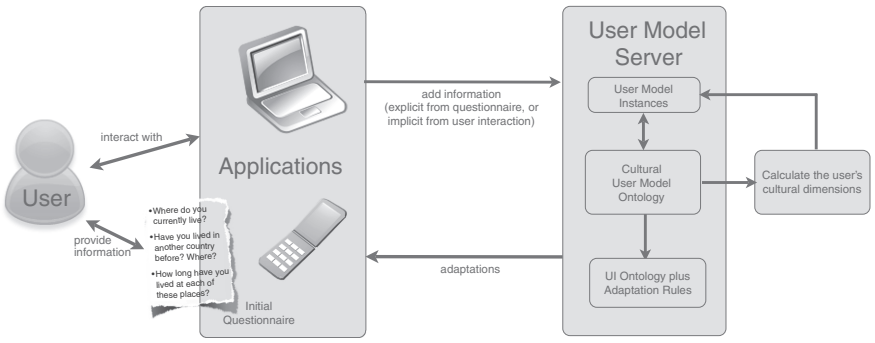


Fig. 1. A framework for cultural adaptivity.

the cultural user model ontology (indicated by the calculation cycle on the right). Specifically, this initial approximation of the user's cultural background includes a weighted average of the durations spent at each country of residence [Reinecke and Bernstein 2008]:

$$influenceOfCountry_N = \frac{monthlyDurationOfStayInCountryN}{ageInMonths}. \quad (1)$$

Note that the sum of all durations of a user's stay in different countries, as represented by the variable *monthlyDurationOfStayInCountryN* is assumed to be roughly equal to the user's age; an estimation that is then used for the variable *ageInMonths* ($\approx \sum_{N \in Countries} monthlyDurationOfStayInCountryN$).

Each country's influence is consecutively multiplied with all cultural dimension scores in order to generate the user's new cultural dimensions. With the help of Hofstede's metrics on five dimensions for each country, we can calculate the user's score in each dimension H (where H is one of Hofstede's 5 dimensions; N the number of countries that influences the user, and *countryScore* is the country's score in the dimensions):

$$userDimScore_H = \sum_{i=1}^{N} countryScore_H * influenceOfCountry_i. \quad (2)$$

Note that the resulting scores are based on the assumption that it is possible to average Hofstede's dimensions for predictive purposes. The approach does not take into account varying intensities of cultural influences at different stages in life. In previous work, this simplification has demonstrated to be robust and suitable to predict preferences for culturally ambiguous users [Reinecke and Bernstein 2008, 2009], although it is certainly not intended to describe a person's cultural background.

After calculating the scores, they are compared to the world averages that are also stored in the cultural user model ontology. The deviations from the world averages can be taken as a basis to assign one of the three classifications low, medium, or high, to each of the user's dimensions. Subsequently, the classification triggers the corresponding adaptation rules, as they are listed in an adaptation ontology (bottom right in Figure 1). As a result, the application triggers adaptations of its interface that suit the user's cultural preferences (bottom arrow).

As an additional possibility in our approach, users can interact with applications and/or devices that are enabled to access the user model server (e.g., an application on a mobile phone, or a computer program). These applications log the user interactions and subsequently inform the user model server about them (top arrow indicating

implicitly added information). Users can also explicitly add or modify information in their personal user model instances. On the server, both interactions and modifications update the user model. This, in turn, triggers adequate adaptations that change the application's user interface.

According to this outline, our requirement for cultural adaptivity is a holistic usability between applications and devices, which could be achieved with a distributed user model, as suggested in Dolog and Nejdl [2003]. Furthermore, the aim is to limit the initial acquisition process to a minimum but still present users with fairly suitable user interfaces before they decide that they do not like the look & feel ("the first impression counts" [Lindgaard et al. 2006]). Since this will not necessarily result in the best possible user interface, a further requirement is the possibility to refine the adaptations both manually by the user, and automatically by the system.

3.3. The Culturally Adaptive System MOCCA

For an evaluation of our approach, we have developed a culturally adaptive system called MOCCA [Reinecke and Bernstein 2009]. MOCCA is a to-do list tool, which helps users to organize their tasks online. It therefore does not provide information itself, but relies on user-generated content. For testing cultural differences in information presentation (as intended by our adaptation rules), this has the advantage that the application does not influence users with culturally-biased content, which could have been the case if we had provided a news application, or similar.

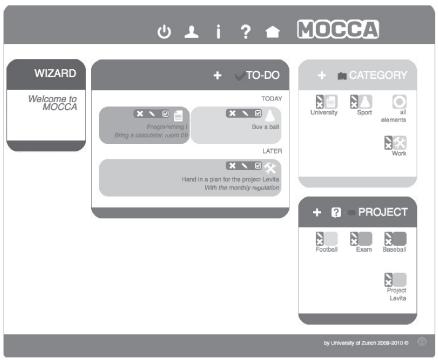
MOCCA can fully adapt its presentation to the cultural background of the user, as it is outlined and stored in the various information units (i.e., the cultural aspects) in the user model. Some example interfaces of MOCCA are shown in Figures 2, 3, and 4.

As specified in our given requirements, the adaptations can occur at different stages of system use, such as after registration, or after observing the user's mouse interactions and inferring the need for refining adaptations. To realize these refinements, MOCCA's user interface needs to be extremely flexible in the composition of different interface components. In particular, each user interface element should be available in different versions. We have therefore tailored the general adaptation rules of Table I to suit the specific application domain of a to-do list application. This resulted in a set of adaptation rules, shown in Table II, that MOCCA uses to trigger adequate modifications of its interface. Note that the left column in Table II lists the adaptable aspects of MOCCA, that is, those interface aspects that have been demonstrated to be influenced by culture (and Hofstede's dimensions).

Offering different user interface elements for each of these adaptable aspects, MOCCA is able to compose a personalized interface. A Chinese user with a score of 80 in the dimension Power Distance, for instance, would be assigned a flat navigation and an icon-represented to-do list similar to Figures 2(a) or 2(b), because the Chinese score is much higher than the world average of 55. Since the dimension Power Distance also relates to the accessibility of functions, MOCCA would also make functionalities, such as edit and delete buttons for to-dos, always accessible (as shown in Figure 2(b)). With that, the user interface automatically becomes more complex. In addition to the information density, the accessibility also changes the interaction possibilities, and thus, common workflows.

In our previous work, MOCCA and its adaptation rules have already demonstrated to be a realistic way of inferring preferences and providing adequate user interfaces for culturally ambiguous users [Reinecke and Bernstein 2008, 2009]. Specifically, our evaluations showed that based on a weighted average of a user's current and former residences, MOCCA is able to provide personalized user interfaces that correspond to the user's preferences by 61% on average (compared to 33% that could be achieved when randomly creating the user interface).

(a) MOCCA with a flat navigation and color-coded to-dos.



(b) An interface with a right-to-left reading direction, a wizard, and constantly visible buttons for deleting, or modifying objects.

(c) MOCCA with little structure, a tree navigation, and a to-do list showing only headers at first sight. The interface shows the dialog to add a new to-do.

(d) MOCCA with a tree navigation, and to-dos that are represented with icons. Question mark bubbles next to the headers offer help-on-demand.

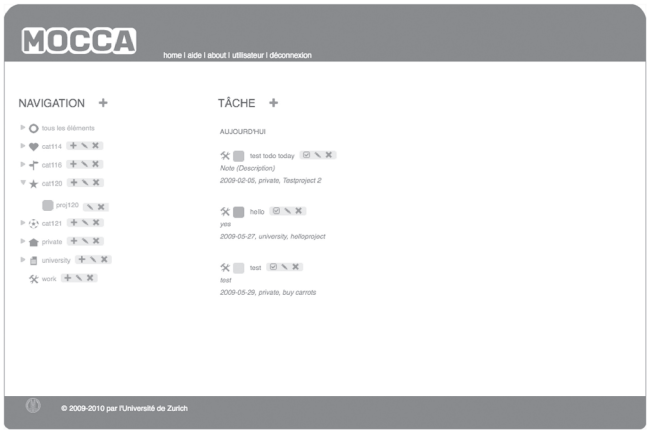
Fig. 2. Example interfaces of the culturally adaptive software MOCCA.

These results are especially promising when considering that very little information can apparently contribute enough to provide users with suitable interfaces right from the start. With that, the approach has demonstrated to be a notable improvement to conventional localization.

We expect that initially asking the user more questions, or deriving information about his or her preferences over time, might further enhance the prediction accuracy. Information about other cultural aspects, such as the user's religion, or education level, for example, could contribute to a more individual user model compared to the initial model of a user's "extended national culture." MOCCA therefore provides two refinement possibilities, as outlined in the approach to cultural adaptivity: (1) a user model editor, in which the user can refine his or her cultural background, and (2) a user interaction tracking mechanism. With that, the system is able to complement the information in the user model by implicitly deriving knowledge from the user's interaction with the interface. Slow mouse movements and much hovering, for instance, indicate that the user is searching for specific information, and MOCCA therefore increases the level of guidance and provides maximum support.

Note that in the following experiment, we concentrate only on the initial acquisition process in accordance with previous research, which had indicated that it is highly important to get the first impression right [Lindgaard and Dudek 2003]. Thus, we

(a) MOCCA's US interface with a hierarchical navigation, which requires clicking on the categories in order to see subordinated projects. Similarly, the to-do list on the right only shows the most important information at first sight. On click, users can expand the information to receive more details.



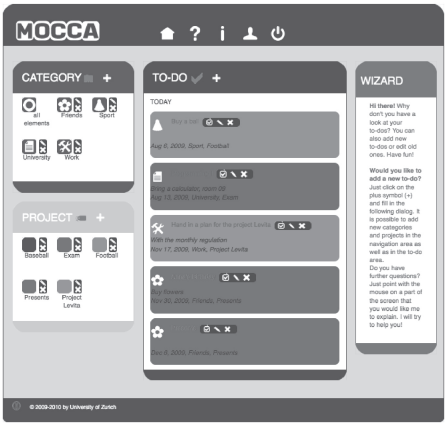
(b) The Swiss version of MOCCA offers the same hierarchical navigation as the US version, where projects can be nested to belong to certain categories. The do-do list showed a medium information density with all information about the to-dos shown at first sight.

Fig. 3. MOCCA's US interface as it was used to represent the benchmark version in the experiment, and MOCCA's Swiss version in comparison (in French)

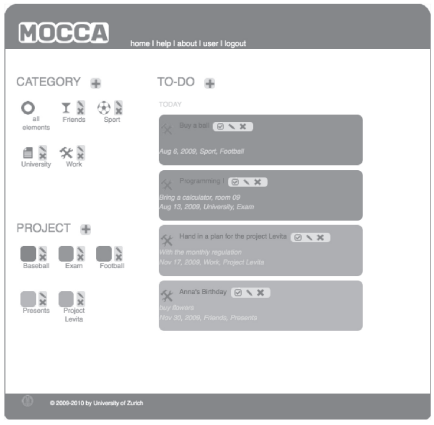
switched off all of MOCCA's rule refinement mechanisms with the effect that the test interfaces were only based on the knowledge about the user's current and former residences, as well as the respective durations.

4. EXPERIMENT ON PERFORMANCE AND USER SATISFACTION

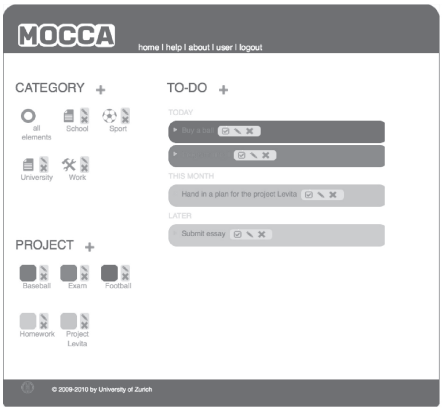
We hypothesize the benefit of cultural adaptivity to be twofold: Firstly, we expect that adapting the interface to a user's "culture" (i.e., his or her current and former



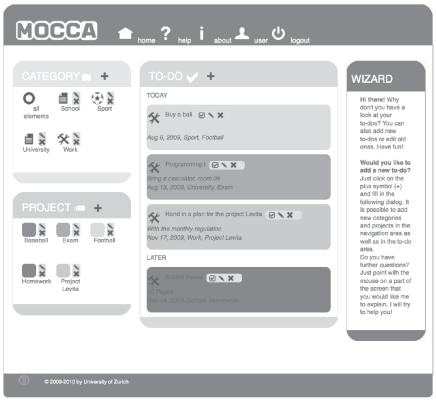
(a) A brightly colored version of MOCCA with a flat button navigation. The participant who received this interface was from Mexico, and had shortly lived in Bulgaria before coming to Switzerland.



(b) A version without the colored borders that define the different areas of the interface. It was personalized for an Indian participant, who had lived in France and the United States for several years, which reduced the complexity of the interface predicted by our model for a "pure" Indian version.



(c) A version with the to-dos in list view, so that to-dos have to be expanded to show more information. The interface was generated for a participant from Poland, who had also lived in Ireland and Germany.



(d) MOCCA with pastel colors, as it was triggered for a participant with Russian, Romanian, and Swiss background. Functions (e.g., delete, add, or edit) are always accessible and add to the information density.

Fig. 4. Example interfaces of MOCCA as they were generated for different participants.

residencies) improves performance (*Hypothesis 1*). If the interface's complexity, its guidance through different dialogs, or its overall support is adapted to the users' needs, it is likely that users will complete tasks and find information in less time and with fewer errors compared to nonadapted systems.

Secondly, we assume that cultural adaptivity increases the aesthetic perception of the interface in comparison to nonadapted interfaces (*Hypothesis 2*). As previously mentioned, it has been suggested that aesthetics might be even more important than the perceived usability [Norman 2004], purely because it is what users become aware of first. In fact, users have been found to rapidly judge the aesthetics of Webpages,

Table II.
MOCCA's adaptable interface aspects and their changes when classified into low, medium, or high. The relation to Hofstede's cultural dimensions is indicated in the second column. This table operationalizes Table I for the to-do list domain

Interface aspect:	Dimension:	Low	Medium	High
Information Density	LTO	To-do items provide minimal information at first sight, requiring the user to click before seeing more information	To-do list shows all information at first sight	Complex version that in addition presents color-encoded information with large icons
Navigation	PDI	Tree menu and to-dos in list view, allows nested sorting	Flat navigation and list view, or tree menu and icon-represented to-do list	Flat navigation and icon-represented to-do list
Accessibility of functions	PDI	Functionalities only appear on mouse-over	Functionalities are always accessible but grayed out if not needed	functionalities are always accessible
Guidance	UAI	When users enter a dialog, all other information in the UI retains visible and accessible	Information other than the current dialog is still visible, but inaccessible	Unnecessary information is hidden in order to force users to concentrate on the currently active dialog
Structure	IDV	Maximum structure: Elements are bordered and affiliations between information is accentuated across elements	Elements are separated and color-coded for better distinction	Minimum structure: Different elements of the UI are only structured through alignment
Colorfulness	IDV	Many different colors	A medium number of colors	The UI is homogeneously colored
Saturation	MAS	Pastel colors with little saturation	Medium saturation and contrast	Highly contrasting, bright colors
Image-to-text ratio	IDV	Image icons in the header menu; category, project, and to-do area hold a representative image	Icons in the header menu are composed of both text and image/s	Header menu consists of textual links only; category, project, and to-do area do not show an image
Support	UAI	On-site support with the help of short tool-tips	The UI offers question mark buttons that expand into help bubbles	An adaptive wizard that is always visible
Help text	PDI	Friendly error messages suggesting how to proceed	Neutral error messages suggesting how to proceed	Strict error messages

PDI = Power Distance Index; IDV = Individualism; MAS = Masculinity;
UAI = Uncertainty Avoidance Index; LTO = Long Term Orientation

often very reliably within 50 milliseconds [Lindgaard et al. 2006]. Tractinsky et al. [2006] validated these results by demonstrating that aesthetic judgements after 500 milliseconds highly correlated with the average attractiveness ratings after an exposure of 10 seconds. However, we know that what someone perceives as beautiful often differs along cultural values. Is cultural adaptivity able to anticipate this divergent perception of beauty?

Moreover, it might not be enough to meet the taste of a person with regards to aesthetics, since 'beautiful' does not mean that users automatically perceive something as usable [Lindgaard and Dudek 2003]. Cultural adaptivity therefore has to generate

interfaces that cater for both the user's aesthetic taste, and usability requirements based on his or her individual cultural background.

In order to evaluate the benefit of cultural adaptivity, our experiment was aimed at comparing the performance and user satisfaction of MOCCA with a personalized version according to a weighted average of the user's current and former residences, compared to MOCCA's US version (see Figure 3(a)). Although none of MOCCA's user interfaces constitute a "null version," the latter has been defined as a benchmark for comparison, since the large majority of software and web sites are still provided by US companies and designers. Thus the following study explores the question, which user interface international users prefer, if comparing the US interface to the culturally adapted interface of MOCCA.

4.1. Method

4.1.1. Participants. We recruited 41 international participants (25 female) from a variety of different cultural backgrounds to avoid restricting the results to only a few national cultures. Participants were between 20 and 38 years old ($m = 26$), all of whom had been living in Switzerland for between 1 and 276 months ($m = 36$ months). Our participants represented 25 different nationalities. We allowed up to 4 people of the same nationality to take part in the study (on average, single nationalities were represented 1.56 times, $sd = .96$). Their former countries of residence, as well as the durations spent in each of these countries were very diverse, with participants having lived in 2-5 different countries previously ($m = 3.1$, $sd = .97$).

Note that for all participants, conventionally localized web sites would provide the Swiss version of their user interfaces (because all of them were living in Switzerland at the time of the user test), which only differs in few aspects to the US version due to a similar cultural classification of both countries (see Figures 3(a) and 3(b)).

In order to balance the education level (and thus, keep this aspect of culture homogeneous), participants had to be students or have completed university (16 had received their master's, 13 held a bachelor's degree, and 11 were currently enrolled in bachelor studies). Participants' study backgrounds were in a variety of fields, ranging from biology to the humanities. However, in order to ensure that they had as little biasing exposure to the experiment as possible, we excluded participants who had taken courses in human-computer interaction, or culture-related topics. This also limited the risk that participants could have consciously or unconsciously anticipated the experiment's objective, or known which version of MOCCA was their personalized one. In addition to this, we controlled for computer literacy. All but one participant were using computers daily (one participant stated she uses the computer a couple of times per week).

Participants were given a small financial incentive for their time.

4.1.2. Procedure. On arrival, participants received both verbal and written explanations about the test procedure, followed by a short questionnaire soliciting information about their cultural background (current and former countries of residence, the durations, their own and their parents' nationality, first and second languages, education level, and religion). In addition, we recorded the frequency of computer usage, age, and gender. Participants were then given a short introduction to MOCCA's purpose and functions, as well as an explanation of its structure of Categories, Projects, and To-Dos. The explanation followed a written script in order to keep it consistent for all participants. Explanations as well as the questionnaire were provided in English.

The test procedure consisted of two subtests testing the US version of MOCCA (our benchmark, see Figure 3(a)) against a personalized version, which MOCCA generated after entering the user's current and former countries of residence into its registration mask (examples of different personalized versions are shown in Figure 4). In order to

Table III. Summary of Evaluation Measurements Used in the Experiment

Usability	• Performance analysis (task completion time, number and type of errors, number of clicks needed to accomplish a task, and number of help requests) • 8-item usability scale on a 7-point Likert scale on effort expectancy, and attitude toward using the system [Venkatesh et al. 2003]
Aesthetics	• 10-item perceived website aesthetic scale on a 7-point Likert scale [Lavie and Tractinsky 2004] • 14-item aesthetic scale with contrary adjectives on a 7-point scale [Hassenzahl et al. 2003; Hassenzahl 2004]
Overall preferences	• 3-items on a 7-point Likert scale on a direct overall, aesthetics, and work efficiency comparison

conceal the US version, this log-in process was performed by the test conductor, who either entered the participant's details, or logged in as a new US participant. We also switched off MOCCA's ability to adapt the language and reading direction, since this could have revealed the interface version. Instead, all participants were presented both interfaces in English and left-to-right alignment.

Each subtest consisted of three equal tasks, which only differed in their wording. Since the tasks partly built on one another, their order remained the same. The first task asked users to create a new category in MOCCA, and subsequently assign this category to a new project. The second one referred to a new to-do, which had to be created following specific instructions, and this had to be placed in the previously constructed project of Task 1. The third task required participants to search for an already existing to-do and its due date by filtering the information on the screen to only show to-dos related to a certain project.

The UI versions (US or adapted) were counterbalanced across participants.

Participants were asked to read through the first set of tasks, which they subsequently had to perform with one version of MOCCA. After each version participants completed the same questionnaire seeking their subjective impression on usability and aesthetics. On completion of the second version and its questionnaire, users were additionally asked to rate the two UI versions in three questions on a 7-item scale, and write down reasons for their preferences. The whole procedure took 30–60 minutes.

4.1.3. Apparatus. We conducted the experiment on an Apple MacBook Pro (2GHz Intel Core Duo, 2GB RAM) with a built-in 15″ LCD display running at the native 1280×800 resolution. Participants had the option of using a keyboard with a Swiss German or US English layout. All participants used an external mouse.

4.1.4. Design and Analysis. The experiment was a within-subjects design with UI version (US, adapted) as the main experimental factor.

Throughout the test, we video-recorded participants to extract the following objective performance measures: time needed for each of the three tasks, number and type of errors, the number of clicks, as well as the number of help requests in MOCCA (see Table III). For later comparisons with these objective results, we also noted participants' verbal reactions. Errors were counted if participants opened a dialogue window that did not lead them to fulfill the tasks; further clicks within the "wrong" dialog counted towards the number of clicks, but not towards the number of errors. In addition, the reported seconds participants needed for each task is the net time, excluding any periods of time spent on explanations or reading.

For the comparisons of time, number of errors, and number of clicks between the two UI versions, we used the nonparametric Wilcoxon signed-rank test for paired samples, since our data was not normally distributed according to Shapiro-Wilk tests ($p < .05$ for all variables). We applied one-tailed significance for testing our directional hypothesis that the adapted version is superior over the US version. All p-values were

corrected for multiple hypothesis testing using the Benjamini-Hochberg correction [Benjamini and Hochberg 1995]. The correction also accounted for the paired samples that were not significantly different.

The nonparametric Friedman's two-way analysis of variance by ranks for related samples was used to check whether the overall distribution of timing data was the same over all three tasks.

Subjective assessments were made for usability and aesthetics with the help of a post-version questionnaire and 7-point Likert scales, where 1 was "I don't agree at all," and 7 was "I completely agree." The comparison of both questionnaires later provided us with an indirect measure of preferences.

For the aspect of usability, the Unified Theory of Acceptance and Use of Technology (UTAUT) [Venkatesh et al. 2003], which in contrast to many other usability tests had been previously validated cross-culturally by Oshlyansky et al. [2007], as well as by Venkatesh and Zhang [2010], was used to collect information about effort expectancy (describing the degree of ease associated with the use of a system), and the attitude toward using the system (describing the user's overall affective reaction to using MOCCA). Initially, we had also included the 4-item scale on self-efficacy (describing the user's perceived competence in mastering the tasks with MOCCA), however, inconsistencies in the answers led us to discard this part of UTAUT.

We additionally used the aesthetics scale of Lavie and Tractinsky [2004], which subdivides an overall impression of aesthetics into classical, and expressive aesthetics. Classical aesthetics are usually referred to as the more traditional notion of design, with factors such as clean, or symmetrical representation of a UI. In contrast, items included in the expressive aesthetics scale (e.g., fascinating, original) aim at capturing the originality or creativity of the design.

We complemented the aesthetic dimensions with Hassenzahl's AttrakDiff [Hassenzahl et al. 2000; Hassenzahl et al. 2003], which directly contrasts the perceived pragmatic quality (i.e., the handling of a product, with variables such as complicated-simple, unpredictable-predictable) with the perceived attractiveness (e.g., quality criteria, such as unpleasant-pleasant, ugly-attractive). Hassenzahl's scale uses bipolar contrasting adjectives as anchors of the scale. We also used this direct comparison of perceived usability with perceived aesthetic quality to investigate a possible *halo effect*, which describes the correlation between two attributes, such as if something is perceived as beautiful it is automatically found to be more usable [Tractinsky et al. 2000]. Such relations were analyzed with Pearson's correlation with a two-tailed test (because we did not have a directional hypothesis that helped us to anticipate whether the relationship between usability and aesthetics would be positive or negative).

For all Likert scale items, we tested their internal consistency using Cronbach's alpha [Cronbach 1951] in order to check for overall reliability, but also to find questions that had been answered in a quite different and inconsistent way. All scales showed high reliability and construct validity with Cronbach alpha scores greater than .744[1] (see Table IV); we therefore computed the averages of participants' responses.

The Likert scale data proved to have significant normal distributions according to Kolmogorov-Smirnov tests on the differences between our dependent scores. Thus, for comparisons of Likert scale data by UI version (US versus adapted), we used dependent t-tests[2] and one-tailed significance in order to test our directional hypothesis.

[1]Cronbachs alpha reliability coefficient ranges between 0 and 1; the closer it is to 1.0 the greater the internal consistency of the items in the scale. For high reliability, Nunally and Bernstein [1994] suggest to use a cut-off of .7.

[2]We therefore assumed that the data can be treated as interval. Non-parametric tests did not change the results.

Table IV. Average Subjective Likert Scale Measures on a 7-point Scale

Likert scale	Rating for US version	Cronbach's alpha	Rating for adapted version	Cronbach's alpha
Effort expectancy	5.77	.830	6.2	.914
Attitude toward using the system	5.01	.915	5.2	.911
Classical Aesthetics	5.52	.851	5.63	.776
Expressive Aesthetics	4.02	.880	4.39	.878
Pragmatic Quality	5.3	.842	5.59	.744
Attractiveness	5.15	.934	5.52	.932

Table V.

Summary of the objective results on performance (+ indicates the better version, and = means that we found no significant difference between the two versions with $\alpha \geq .1$). P-values have been adjusted for lower significance (thus higher p-values), using the Benjamini-Hochberg adjustment for multiple hypothesis testing, including the paired samples that were not significantly different

Measures	Task	US version	adapted version	p-value
Task completion time	All	−	+	$p < .05$
	1	−	+	$p < .05$
	2	=	=	n. s.
	3	−	+	$p < .05$
Number of clicks	All	−	+	$p < .001$
	1	−	+	$p < .05$
	2	−	+	$p < .05$
	3	−	+	$p < .01$
Error rate	All	−	+	$p < .001$
	1	−	+	$p < .01$
	2	−	+	$p < .05$
	3	−	+	$p < .01$

All p-values were again corrected for multiple hypothesis testing with the Benjamini-Hochberg correction [Benjamini and Hochberg 1995].

Possible interaction effects of the UI version on aesthetics and user experience were analyzed with a repeated-measure ANOVA with the dimension (e.g., classical, expressive) and UI version (US, adapted) as within-subject factors.

The experiment ended with three questions on the participants' overall preferences, which directly compared the two UI versions on a 7-point scale (1 = the first version, 4 = neutral, 7 = the second version; later converted to 1 = US version, 4 = neutral, 7 = adapted version). Participants had to answer which version they liked best, which one they found more aesthetically appealing, and which one they could work with more effectively. Correlations between these overall answers and the previously recorded perceived usability and aesthetics were again investigated with the help of a Pearson correlation and a two-tailed test. Furthermore, we tested whether a significant majority preferred one version over the other with the chi-square goodness of fit test, entering the preferred version as a categorical variable with the three levels preferredAdapted, preferredUS, and neutral.

4.2. Results

An overview of our results on the objective performance measures is provided in Table V, and the subjective results are presented in Table VI.

4.2.1. Performance. The distribution of timing data was significantly different for all three tasks and UI versions (Friedman's two-way ANOVA, $\chi^2_{(5)} = 42.03$, $p < .001$). We used post-hoc Wilcoxon tests to follow up this finding.

The overall difference in time needed to complete all three tasks proved a notable advantage for the adapted version ($Z = -2.002$, $p < .05$, $r = -.22$) with participants

Table VI.
Summary of the subjective results on the US version versus the adapted version of MOCCA (+ indicates the better version, − worse, +/= and −/= describes a trend observed at a confidence level of $\alpha < .1$), and = means that we found no significant difference between the two versions with $\alpha \geq .1$). P-values have been corrected with the Benjamini-Hochberg adjustment for multiple hypothesis testing, including the paired samples that were not significantly different

Measures		US version	Adapted version	p-value
Usability	Effort expectancy	−	+	$p < .05$
	Attitude toward using the system	=	=	n. s.
Aesthetics	Classical aesthetics	=	=	n. s.
	Expressive aesthetics	−	+	$p < .05$
User Experience	Pragmatic quality	−/=	+/=	$p < .1$
	Attractiveness	−	+	$p < .05$
Overall preferences	Overall preferred	−	+	−
	Aesthetically preferred	−	+	−
	Work efficiency preferred	−	+	−

taking 276.46 seconds on average to complete all tasks with the US version ($sd = 129.9$), versus 215.39 seconds ($sd = 98.6$) with the adapted version. This equals an average time improvement of 22%.

Participants needed on average 92.37 seconds to complete Task 1 with the adapted version ($sd = 61.2$), but 120.98 seconds ($sd = 75.05$) with the US version, indicating an improved efficiency when working with a culturally adapted interface ($Z = -1.87$, $p < .05, r = -.21$).

Task 2 was also on average performed faster with the adapted version ($m = 71.29$ seconds, $sd = 25.4$) than with the US version ($m = 83.51$ seconds, $sd = 53.79$), though not significantly.

Task 3 asked participants to find a given to-do and write down its due date. The task was typically completed in less time than the other two, with the fastest participant accomplishing it within 11 seconds using the adapted version. In general, the completion took significantly more time with the US version ($m = 71.09$ seconds, $sd = 44.08$) than with the adapted version ($m = 51.73$ seconds, $sd = 29.2$), $Z = -2.12$, $p < .05$, $r = -.23$.

The time needed for each task was also mirrored in the number of clicks, demonstrating a significant advantage for the adapted version ($Z = -3.40$, $p < .001$. $r = -.38$). Participants typically needed 22 % more clicks when using the US version.

For task 1, participants used 13.9 clicks ($sd = 6.86$) on average for the US version, whereas the adapted interface significantly lowered this number to 11.68 ($sd = 4.38$), $Z = -2.06$, $p < .05$, $r = -.23$. The same trend was observed for tasks 2 and 3: Task 2 was accomplished with 9.32 clicks on average for the adapted version ($sd = 2.1$) versus 11.59 clicks ($sd = 6.12$) for the US version (significantly more with $Z = -2.11$, $p < .05$, $r = -.23$). Task 3 could be accomplished with only one click (achieved by 1 participant using the US version, and 12 participants using the adapted version). However, on average participants needed 5.27 clicks ($sd = 3.76$) to accomplish the task with the US version, but only 2.85 clicks ($sd = 2.14$) in the adapted version ($Z = -3.39$, $p < .01$, $r = -.37$).

Performance was additionally measured by recording the number of errors. Participants made 69% fewer errors with the adapted version than with the US interface ($Z = -3.73$, $p < .001, r = -.53$).

Naturally, participants made the most errors during task 1 when still getting to know the user interface. With an average error rate of 1.27 ($sd = 1.42$), however, the US version caused significantly more errors for this task than the adapted version

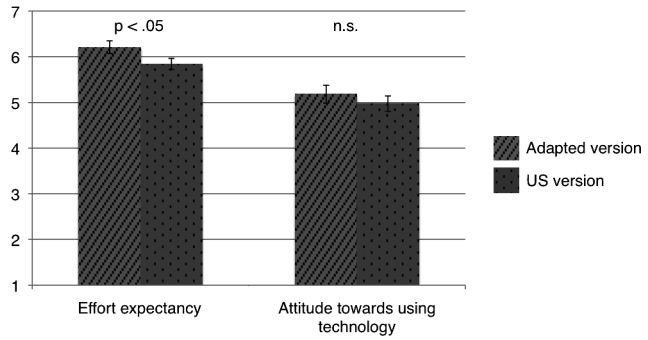


Fig. 5. Average evaluation scores of UTAUT's effort expectancy and attitude toward using the system for MOCCA's US version and the adapted user interface. Error bars represent the standard error.

($m = .51, sd = 1.08$; $Z = -2.8$, $p < .01$, $r = -.31$). Two participants also requested the system's help whilst using the US version. Both of them were using this version first, therefore this cannot be rated negatively (although there were no help requests while using the adapted version). Task 2 also showed a lower error rate for the adapted version ($m = .02$, $sd = .16$), than for the US version ($m = .24$, $sd = .62$; $Z = -2.12$, $p < .05, r = -.23$), and this advantage for the adapted version was also shown for task 3 (adapted version: $m = .20$, $sd = .56$, versus $m = .88, sd = 1.25$ for the US version; $Z = -3.07$, $p < .01, r = -.34$).

4.2.2. Usability. Subjective usability results are shown in Figure 5. The effort expectancy (e.g., "My interaction with MOCCA is clear and understandable," "I find MOCCA easy to use") was perceived significantly higher for the adapted version ($m = 6.2, sd = .77$) than for the US version ($m = 5.77, sd = .88$; $t_{(40)} = -2.46$, $p < .05$). We did not find significant differences in the attitude toward using the system (e.g. "Working with MOCCA is fun", "MOCCA makes organizing to-dos more interesting") between the two versions, although the adapted version was again rated slightly better.

4.2.3. Aesthetics. The results for classical aesthetics did not show a strong tendency towards one UI version, with similar average ratings for both (5.52 for the US version, and a slightly higher 5.63 for the adapted version). Thus, both versions seem to satisfy participants' traditional aesthetics sensibility.

The expressive aesthetics received an overall lower rating than the classical aesthetics, but the adapted version was judged significantly better ($m = 4.39, sd = 1.25$) than the US version ($m = 4.02, sd = 1.35$; $t_{(40)} = -2.17$, $p < .05$).

The average scores for the two aesthetic factors as a function of the factor UI version are shown in Figure 6.

To investigate the interaction between UI version and aesthetic factors, we ran a 2×2 repeated-measures ANOVA, which showed a significant effect for the aesthetic factors classical and expressive ($F_{(1,40)} = 95.33$, $p < .001$), but not for the UI version.

An analysis of the interaction between UI version and aesthetic scale further indicated that the effect of the UI version on aesthetics is slightly modulated by the aesthetics dimension, with a statistical significance at the 90% level ($F_{(1,40)} = 3.28$, $p = .078$).

4.2.4. User Experience. The user experience measures were meant to combine participants' impressions on usability and aesthetics, and the results (see Figure 7) verified the above-mentioned tendencies towards a preference for the adapted version. The pragmatic scale, evaluating the perceived ease of handling of MOCCA, resulted in an average rating of 5.3 ($sd = .92$) for the US version, and an average of 5.59 ($sd = .66$)

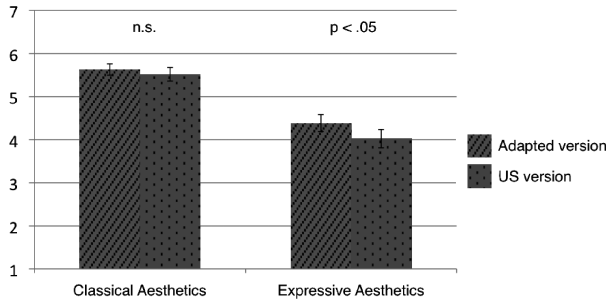


Fig. 6. Average evaluation scores of classical and expressive aesthetics for MOCCA's US version and the adapted user interface. Users especially favored the expressive aesthetics of the adapted version over the US version, showing that they found the design more original and creative. Error bars represent the standard error.

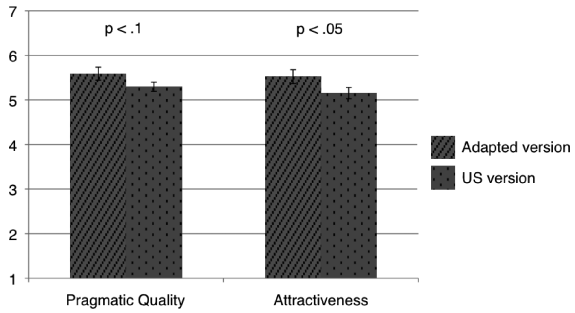


Fig. 7. Average evaluation scores for the pragmatic quality and attractiveness with UI version as a factor. Error bars represent the standard error.

for the adapted version, showing a slight tendency towards an improved perceived handling of the adapted version ($t_{(40)} = -1.799$, $p = .06$).

The attractiveness, often considered as an equal contributor to the overall observed usability [Tractinsky et al. 2000], scored significantly higher for the adapted version ($m = 5.52$, $sd = .82$) than for the US version ($m = 5.15$, $sd = 1$), $t_{(40)} = -2.76$, $p = .05$.

In addition, Pearson's correlation was significant between the pragmatic quality and attractiveness for both versions (US: $\rho_{(41)} = .79$, $p < .001$, adapted: $\rho_{(41)} = .62$, $p < .001$), indicating a possible halo effect between the perceived aesthetics and the perceived ease of handling. We found a significant effect for the UI version ($F_{(1,40)} = 6.35$, $p < .05$), but not for the user experience measures.

4.2.5. Overall Preferences and Qualitative Feedback. Comparing the two versions at the end, the majority of participants preferred the adapted version, and this preference was especially strong for the questions "Which version did you like best?", and "Which version did you find more aesthetically appealing?" (see Figure 8). The distribution of answers to all three questions was skewed towards the adapted version:

For the first question, 51% of the participants rated on the extreme ends of the scale (i.e. 1 = US version, or 7 = adapted version), with 34% of the participants strongly favoring the adapted version, versus 17% who preferred the US version. On the 7-point scale, this resulted in an average rating of 4.76 ($sd = 2.34$). In order to include the tendencies towards one version, we subdivided the scale into two parts (i.e., 1–3 and 5–7). Combining the choices of each subscale, 66% of the participants preferred the adapted version ($m = 1.7$ on a subscale of 1-3, $sd = .82$), 29% preferred the US version

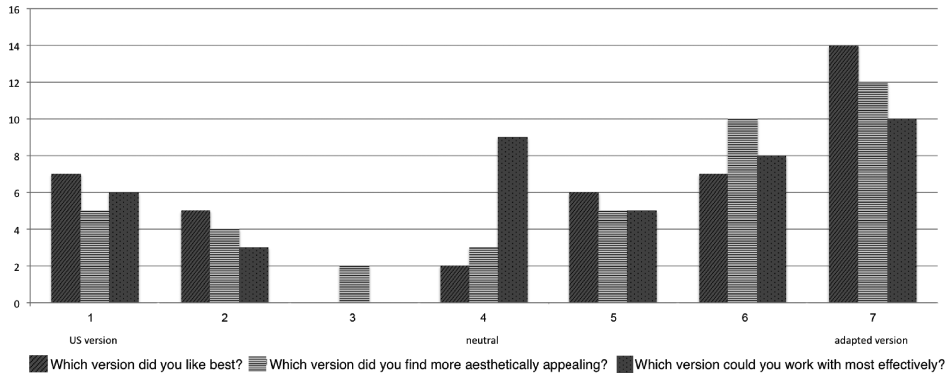


Fig. 8. Raw counts of participants' responses to the three preference questions on three 7-point Likert scales.

($m = 1.42$, $sd = .51$), and 5% were neutral. This advantage for the adapted version was highly significant $\left(\chi^2_{(2)} = 23.17, p < .001\right)$.

The second question on which version participants found more attractive showed a similar trend: 29% of the participants favored the adapted version and marked the 7 on the scale, versus 12% who were clearly in favor of the US version (marking 1). The average rating was 4.88 ($sd = 2.15$). On the two sub-scales, 66% of the participants found the adapted version more aesthetically appealing (85% of them were the same participants who had also chosen the adapted version as the overall preferred one). In contrast, 27% preferred the aesthetics of the US version, and 7% were neutral. The results show that the majority of participants classified the attractiveness of the adapted version much higher $\left(\chi^2_{(2)} = 21.85, p < .001\right)$.

Participants did not show such clear preferences towards the version they could work with most effectively: 24% highly preferred the adapted version (i.e. rated with 7), while 15% opted for the US version (rated with 1). However, a relatively high proportion at 22% were neutral towards both versions. Nevertheless, the distribution of answers was skewed towards the adapted version, with an average rating of 4.66 ($sd = 2.09$). Combining the answers on each end of the scale, 56% perceived the work efficiency as better with the adapted version, versus 22% with the US version, and this preference was significant $\left(\chi^2_{(2)} = 9.56, p < .01\right)$.

Interestingly, the objective performance results partly contradict the subjective feelings of those participants who had stated that they could work better with the US version, because their timing data was not significantly different to the time needed for the adapted version across all three tasks (US: 80.0s, sd = 52.27s; adapted: 74.4s, sd = 33.25s). In contrast, participants who thought they could work more efficiently with the adapted version, did need less time with this version for all three tasks (US: 95.8s, sd = 67.5; adapted: 74.12s, sd = 51.22; $Z = -2$, $p < .05$, $r = -.22$).

Participants' responses to the question, which version they liked best, correlated slightly higher with their opinion on which version they could work with most efficiently ($\rho_{(41)} = .875$, $p < .001$), than with their aesthetic preferences ($\rho_{(41)} = .558$, $p < .001$). This could imply that the perceived work efficiency is the crucial factor when being forced to decide between two user interfaces. We therefore analyzed whether participants' written explanations for their preferences confirm this idea. Specifically, we used participants' written comments on the positive aspects of each version to find possible explanations for their choice for one version or the other. The positive keywords mentioned during this final explanation part are listed in Table VII.

Table VII.
Overview of keywords in participant's written responses to the question why they preferred one version over the other.)

	Comment	US version	Adapted version
Usability	easier to use	3	8
	guidance through dialogs	–	1
	clear	4	6
	practical	–	1
	overview of to-dos	1	3
	simplicity	10	1
	support	–	2
	flat/hierarchical navigation	2	1
	intuitive	–	1
	structure/organization of the UI	1	7
	more predictable	–	1
Total		**21**	**32**
Aesthetics	icons	1	5
	motivating	–	1
	appealing	2	5
	colors	1	13
	creative	–	2
	inviting	–	1
	pleasure/fun	–	2
	formal/informal	4	2
Total		**8**	**31**

Altogether, usability-related aspects were mentioned 53 times, whereas the aesthetics of the user interface was referred to 31 times. Thus, most participants seemed to base their comparison not on high-level features, such as the number or kind of colors, or the complexity of the user interface, but focused on practical handling aspects. The comments therefore further substantiate the assumption that work efficiency and effectiveness could be the more important factors.

A positive aspect of the US version that was mentioned most often was its simplicity (mentioned 10 times). Four participants who had acknowledged this found the colors of the adapted version nicer, however, for them, the simplicity of the US version seemed to be the decisive factor leading to an overall preference for the US version. Again, this finding further verified that participants' overall preferences were mainly based on aspects of usability.

For the adapted version, its ease of use was positively acknowledged the most (8 times). Those participants whose adapted version included a clear subdivision of categories, projects, and to-dos with the help of colored borders (as shown in Figures 4(a) or 4(d)) approved of this structure (mentioned 7 times). Only 1 participant acknowledged the increased guidance through dialogues in her adapted version, and 2 participants gave more support (i.e., help bubbles or a wizard) as reasons for their preference for the adapted version. Altogether, the usability of the adapted version was positively commented on 32 times, versus 21 comments for the US version, a result that is consistent with the overall majority preference for the adapted version.

Aesthetic aspects were positively acknowledged for the US version by 8 participants, and by 31 participants for the adapted version. Most participants commented on the colors, which were preferred by 13 in the adapted version. It has to be noted that MOCCA's US interface uses very few, monotonous colors due to the US individualism score of 91, which is the highest of all countries that Hofstede compared. The average participant in our test received a score of 48, $sd = 18.04$, and only 17 participants were presented with an equally monotonous color scheme as the US version (although mostly consisting of a different color selection because of the influence of the Masculinity score on the

saturation of colors). Most participants were therefore presented with more colorful interfaces; 54.5% of these participants mentioned the colors as a positive aspect, showing that this is one of the most striking characteristics in MOCCA. This also emphasizes the importance of changing the color schemes when adapting to different cultural backgrounds.

Interestingly, 4 participants found the look of the US version very formal, and said it seemed to be designed for work purposes. Two of them also compared it to the adapted version, which they thought looks as if it was designed for leisure activities. The comments make it clear that preferences often depend on the context of Websites. Since MOCCA's purpose is to support users' planning activities, some of our participants might have found the design of the US version more appropriate. It is even more surprising, that the majority still preferred the adapted interface.

Furthermore, participants' comments were helpful to interpret why some users with divergent opinions on their aesthetics and work efficiency preferences, had a very specific tendency towards one version when asked for their favorite user interface. For example, one participant with a Japanese background preferred the color and layout of the adapted version, and he liked that the divisions (category, project, to-do) are explicitly shown on the screen: "The [US version] is better to understand, though the [adapted version] is better in design." In the overall comparison ('Which version did you like best?"), however, he marked the second box on a 1 (= US version) to 7 (= adapted version) scale. For the question on which version offered him the best work efficiency, he even ticked the first point, indicating the highest preference for the US version. One explanation for his contradictory answers might be that he used the adapted version first, and probably included his work efficiency improvements into the final judgements. In contrast, we were generally not able to find a correlation between the first version participants used, and his or her final overall preference: 14 participants who had used the adapted version first, and thus, experienced an equally steep learning curve at the start, still preferred this in the overall preference rating (only 6 did so for the US version). Of those participants, who had preferred the adapted version in the end, 13 participants had used the US variant first, versus 6 who later preferred the US one.

Surprisingly, we observed the opposite effect (an overall preference for the adapted version, but a strong preference for the aesthetics of the US version) in only one case, where a participant with Indian background stated that he liked the US version for its simplicity, but found the adapted version more appealing, more creative, and more innovative. For him, the simplicity seemed to be more appropriate for a to-do application, which is why he preferred the look of the US version. Again, this corresponds to previous findings indicating that preferences depend on the domain.

Two participants also justified their preference for the US version by saying that it reminded them of the social network site Facebook, which uses similar colors.

Although 21 participants rated their overall preference on the 'extreme' ends of the scale (7 voted 1, i.e., the US version, and 14 voted 7, i.e., the adapted version), only 5 of them affirmed that they would refrain from using MOCCA if presented the other version, all of whom referred to "ugly colors" as the main reason.

4.2.6. Possible Influences on the Users' Preferences. As our results showed a marked preference for the adapted version by most participants, we have to consider the possibility that the US version we presented is simply unattractive, or flawed in some other respect. There are two main arguments against this: Firstly, the ratings for the classical aesthetics were similar for both interfaces presented, and secondly, the US version did not receive extraordinarily low ratings, indicating that users did appreciate its usability.

Table VIII.
Correlations between perceived pragmatic quality and perceived attractiveness for Westerners, and Latin American/Asian participants (analyzed with Pearson's correlation and a two-tailed test, $*p < .05$, $**\ p < .01$, $***p < .001$)

	Westerners	non-Westerners (Asians and Latin Americans)
US version	$\rho_{(22)} = .81***$	$\rho_{(19)} = .79***$
Adapted version	$\rho_{(22)} = .64**$	$\rho_{(19)} = .49*$

According to our approach to cultural adaptivity, we expected MOCCA's US version to be rated higher by participants with a cultural background closer to the US ('Westerners'). These participants were thought to be less decisive in their preference for the adapted version than those participants who had lived in Asian or Latin American countries for most of their lives. The latter are considered to be collectivist countries (receiving a low score in the dimension Individualism), who have been found to prefer colorful interfaces, where color is also used to show affiliations between information, and structure the interface (cf. the adaptation rules in Table II). Hence, MOCCA's user interfaces generated for Asians and Latin Americans were much more colorful, than most of those ones generated for Westerners.

To analyze whether Westerners chose the US version over their personalized interface more often than Latin American and Asian users, we subdivided the user population into two groups according to the country where they had spent most of their lives. This resulted in 19 people who were assigned to the Asian and Latin American cluster, and 22 participants to the Western cluster.

The answers for participants' overall preferences revealed that within both groups a significant majority of users preferred the adapted version over the US interface (non-Westerners: $\chi^2_{(1)} = 12.96$, $p < .001$, Westerners: $\chi^2_{(1)} = 23.04$, $p < .001$). However, this number was lower for Westerners (68%, compared to 74% of the non-Westerners), supporting our assumption that Westerners, who are closer to the US culture than Latin Americans or Asians, were less decisive in their preference for the adapted version.

We also found a significant relationship between the perceived pragmatic quality and the perceived attractiveness for both groups as shown in Table VIII. The observation suggests that the better participants rated the pragmatic quality of MOCCA's interfaces, the better they classified their attractiveness. The correlation between the two variables is not significantly different for Westerners and non-Westerners (using Fisher r-to-z-transformation), suggesting that the finding is independent of cultural affiliation.

Furthermore, we had assumed that the (usually more colorful) adapted versions would be mostly rejected by male participants. Instead, of the 12 participants who preferred the US version, only 3 were male. Two participants, who preferred the US version, had received a personalized interface in our evaluation that was fairly close to the US interface at first sight due to similar cultural dimensions: the only visible differences were an expanded to-do list, and an increased support with question mark bubbles which opened upon click. The color scheme, and the hierarchical navigation, however, were equivalent to the US version. The remaining 10 participants were of divergent cultural backgrounds, which resulted in mostly colorful interfaces with a flat navigation, an emphasis on the structure and affiliation between to-dos and the categories/projects they belonged to. These participants gave either a preference for few colors, or the simplicity of the US version as a reason for their preference.

4.3. Discussion

Our participants were significantly faster with the adapted version, needed fewer clicks to complete tasks, and made fewer errors. The results were supported by their perceived effort expectancy, which was significantly better for the adapted version.

While we were not able to find a difference for the ratings of classical aesthetics between both versions, this suggests that MOCCA's interface design satisfies traditional design perceptions (symmetric, clean, pleasant) no matter which version is shown. In contrast, the expressive aesthetics, which stand for more original and creative design, were rated significantly higher for the adapted version. An advantage for the attractiveness (describing the user experience) of the adapted version supported this result.

Participants' answers in a direct comparison of the two versions were particularly strong: 56% thought they could work more efficiently with the adapted version, 66% found the adapted version more aesthetically appealing, and 66% opted for the adapted version in an overall comparison. These answers to the last three questions asked for a direct comparison between the two versions, and this may have lead many participants to subconsciously perceive the 7-point scale as a dichotomous choice. In a real-life situation, participants would indeed have to decide between one version and the other (in case there is an alternative to a certain Webpage); our results suggest that the adapted version of MOCCA would outperform competing nonadapted Websites.

Moreover, our findings indicate that cultural adaptivity has an advantage over localized web sites. Conventional approaches to localization would have (more or less automatically) presented users with the Swiss version, because this is where all our participants lived at the time of the experiment. MOCCA's Swiss version (localized), however, is very similar to the US version (nonlocalized); it is therefore likely that the majority of our participants would favor the adapted interface over the Swiss variant.

Our results demonstrate that the perceived work efficiency outweighs the perceived aesthetics when participants have to decide between a nonadapted, and an adapted user interface. In particular, users' responses to why they chose one version over the other suggested that usability was a more crucial factor in their decision than aspects related to the attractiveness. The finding contradicts the assumption of Ben-Bassat et al. [2006], who had speculated that in lab experiments participants give aesthetics a stronger weighting than usability-related aspects, because a less usable system does not carry any consequences. Our results of the subjective ratings of perceived usability and aesthetics indicate that the ratings on these two factors significantly relate to one another. This relationship was also strong if analyzing the ratings for Westerners and non-Westerners (Asians and Latin Americans) separately, suggesting that a halo effect between pragmatic quality and attractiveness is culturally independent. Tractinsky [1997] had previously observed a similar relationship between perceived aesthetics and a priori perceived usability, when doing a study with Japanese and Israeli subjects. While he assumed a possible cultural difference in the magnitude of the relationship (with Israelis perceiving the two factors as more related than Japanese participants), our findings did not support this when dividing participants into Westerners and non-Westerners.

5. LIMITATIONS AND FUTURE WORK

Our results showed a strong preference for the culturally personalized version of MOCCA, if providing participants with a choice between this and the US version. Due to our experimental design, which aimed to investigate this superiority for a broad range of cultures (i.e., culturally ambiguous users), we are able to say that cultural adaptivity is able to outperform nonadapted versions of the same Website; however, we cannot claim that one of MOCCA's personalized sites is the best suitable for a user. Instead, it is well possible that our participants preferred the adapted version, but would have also said so if they were presented the adapted interface of another participant. To investigate the generalizability of the results, we therefore need to conduct large-scale studies in the future, where participants can be presented with several different versions of one web site.

Future experiments should also evaluate the users' perception of the adapted version versus a nonadapted interface at different stages of usage. So far, our study strongly indicates an advantage for the culturally personalized interface, and a closer look at participants' responses suggested that the perceived usability was the determining factor for this preference. However, the first impression of Webpages usually influences the perceived aesthetics, rather than assumptions on whether the site will be easy to use. Our study was not designed to answer the question which factor would be more pivotal to leaving a Website for the competition, but rather whether users were more inclined to do so when using MOCCA's nonadapted version, which our results did suggest. Evaluating this finding in detail will be an interesting goal in the future.

Similarly, an exciting direction for the future will be to compare our results with those of a long-term study. We can see several open questions, which require more extensive studies:

First, it will be interesting to evaluate whether the results remain similar after participants have used MOCCA for a while. In particular, we expect the performance improvement of the adapted version to slowly equate that of the US version after participants have become more acquainted with the system.

Second, a long-term study can provide us with some insight on whether users will adopt their personalized site, or would prefer to change the look & feel, if provided with the ability to manually modify certain parts of the interface.

Third, we will evaluate whether culturally adaptive user interfaces retain the recognition value. In MOCCA, we have tried to support the brand recognition with the help of a logo at the top of the page, and enframing the site with a header and a footer that have a similar appearance over different interface versions. So far it is unclear whether this is enough to maintain the effect of remembering a brand. We could imagine that companies might want to reduce the adaptation possibilities, so that, for instance, at least the main colors remain constant to the company's branding, and this would be very well possible with our approach. Yet further studies should investigate which adaptation rules are most important for increasing the users' satisfaction; if color is one of the main factors determining users' satisfaction with a site (as suggested by the results of Lindgaard [2007] as well as our own), then excluding the automatic adaptation of colors will impair the success of cultural adaptivity.

6. CONCLUSION

We evaluated our culturally adaptive system MOCCA, which is able to provide users with personalized interfaces according to a weighted average of the user's duration of stay at current and former residences. Our results substantiate the idea that preferences differ, and show that culturally adaptive user interfaces can reflect these differences. In particular, we have demonstrated that the method has a competitive advantage over nonadapted user interfaces. First of all, our experiment showed that users' performance metrics significantly increased by 22–69%, supporting our first hypothesis. The result was emphasized by users' perceived usability, showing that they considered the adapted version to be significantly easier to use. In line with our second hypothesis, users were also more satisfied with the personalized user interface when asked about their perception of the attractiveness.

In an overall comparison of the two interface, all of the aforementioned results were again verified: a significant majority of 66% favored MOCCA's culturally adapted interface, a remarkable 66% also found it more aesthetically appealing, and 56% thought they could work more efficiently with this version.

These results demonstrate an exceptional benefit for culturally adapted interfaces over providing users with a Webpage's "standard" version. They also indicate that the conventional understanding of "good" user interface design has to be seen in the context

of cultural differences: In our eyes, it is not feasible to find a magic formula for what international users perceive as usable and beautiful, and correspondingly, the practice of designing one interface for all is unlikely to satisfy users' expectations. In contrast, culturally adaptive interfaces seem to be a promising solution to anticipate what users like, and to improve their user experience, no matter where they come from.

Summarizing all the findings, we found strong evidence that cultural adaptation of user interfaces plays a central role in ensuring both the work efficiency and user interface acceptance in the globalized world of the IT industry. The cultural adaptivity framework and our MOCCA application are good examples of how to technically support this increase in user satisfaction and task performance. We hope that our investigation will encourage others to explore and expand on this important element of usability to overcome the cultural divide.

ACKNOWLEDGMENTS

The authors would like to thank Anthony Jameson and Robinson Aschoff for numerous valuable discussions on the test design and the results, as well as the anonymous reviewers and the editor for many helpful comments.

REFERENCES

ACKERMAN, S. 2002. Mapping user interface design to culture dimensions. In *Proceedings of the International Workshop on Internationalization of Products and Systems (IWIPS'02)*.

BADRE, A. 2000. The effects of cross cultural interface design orientation on World Wide Web user performance. GVU Tech. rep.

BARBER, W. AND BADRE, A. 1998. Culturability: The merging of culture and usability. In *Proceedings of the Conference on Human Factors and the Web.*

BAUMGARTNER, V.-J. 2003. A practical set of cultural dimensions for global user-interface analysis and design. M.S. thesis, Fachhochschule Joanneum, Austria.

BEN-BASSAT, T., MEYER, J., AND TRACTINSKY, N. 2006. Economic and subjective measures of the perceived value of aesthetics and usability. *ACM Trans. Comput.-Hum. Interact. 13*, 2, 210–234.

BENJAMINI, Y. AND HOCHBERG, Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Royal Stat. Soc. Series B Method. 57*, 1.

BENYON, D. 1993. Adaptive systems: A Solution to usability problems. *User Model. User-Adapt. Interact. 3*, 65–87.

BLOCH, P. 1995. Seeking the ideal form: product design and consumer response. *J. Market. 59*, 3, 16–29.

BOURGES-WALDEGG, P. AND SCRIVENER, A. 1998. Meaning: the central issue in cross-cultural HCI design. *Interact. Comput. 9*, 3, 287–309.

BURGMANN, 1., KITCHEN, P., AND WILLIAMS, R. 2006. Does culture matter on the Web? *Market. Intel. Plan. 24*, 1,62–73.

CALLAHAN, E. 2005. Cultural similarities and differences in the design of university Websites. *J. Comput.-Mediat. Comm. 11*, 1.

CHA, H., OSHLYANSKY, L., AND CAIRNS, P. 2005. Mobile phone preferences and values: The U.K. vs. Korea. In *Proceedings of the International Workshop on Internationalization of Products and Systems (IWIPS)*.

CHOI, B., LEE, I., KIM, J., AND JEON, Y. 2005. A qualitative cross-national study of cultural influences on mobile data service design. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI'05)*. 661–670.

CORBITT, B., THANASANKIT, T., AND HAYNES, J. 2002. A model for culturally informed web interfaces. In *Internet Management Issues: A Global Perspective*, IGI Global, 1–26.

CRONBACH, L. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika 16*, 297–334.

DOLOG, P. AND NEJDL, W. 2003. Personalisation in Elena: How to cope with personalisation in distributed eLearning networks. In *Proceedings of the International Conference on Worldwide Coherent Workforce, Satisfied Users—New Services For Scientific Information*.

DORMANN, C. AND CHISALITA, C. 2002. Cultural values in Web site design. In *Proceedings of the European Conference on Cognitive Ergonomics.*

FINDLATER, L. AND MCGRENERE, J. 2008. Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI'08)*. 1247–1256.

FINDLATER, L., MOFFATT, K., MCGRENERE, J., AND DAWSON, J. 2009. Ephemeral adaptation: The use of gradual onset to improve menu selection performance. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI'09)*. 1655–1664.

FORD, G. AND GELDERBLOM, H. 2003. The effects of culture on performance achieved through the use of human computer interaction. In *Proceedings of the Conference on Enablement Through Technology.*

GAJOS, K., WOBBROCK, J., AND WELD, D. S. 2008. Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI'08)*.

GOULD, E., ZAKARIA, N., AND YUSOF, S. 2000. Applying culture to Website design: A comparison of Malaysian and U.S. Websites. In *Proceedings of theJoint IEEE International and 18th Annual Conference on Computer Documentation*. 161–171.

GREENBERG, S. AND WITTEN, I. 1985. Adaptive personalized interfaces: A question of viability. *Behav. Inform. Techn. 4*, 31–45.

GUPTA, A. AND FERGUSON, J. 1997. *Anthropological Locations: Boundaries and Grounds of a Field Science.* University of California Press, Berkeley, CA.

HASSENZAHL, M. 2004. The interplay of beauty, goodness, and usability in interactive products. *Hum.-Comput. Interact. 19*, 4, 319–349.

HASSENZAHL, M., BURMESTER, M., AND KOLLER, F. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualitat. In *Mensch & Computer 2003: Interaktion in Bewegung*, 187–196.

HASSENZAHL, M., PLATZ, A., BURMESTER, M., AND LEHNER, K. 2000. Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI'00)*. 201–208.

HEIMGÄRTNER, R. 2005. Towards cross-cultural adaptive human-machine-interaction in automotive navigation systems. In *Proceedings of the 7th International Workshop on Internationalisation of Products and Systems*. 97–111.

HEIMGÄRTNER, R., TIEDE, L.-W., LEIMBACH, J., ZEHNER, S., AND WINDL, H. 2007. Towards cultural adaptability to broaden universal access in future interfaces of driver information systems. In *Proceedings of the 4th International Conference on Universal Access in Human-Computer Interaction: Ambient Interaction.* 383–392.

HERMEKING, M. 2005. Culture and Internet consumption: Contributions from cross-cultural marketing and advertising research. *J. Comput.-Mediat. Comm. 11*, 1.

HODEMACHER, D., JARMAN, F., AND MANDL, T. 2005. Kultur und Web-Design: Ein empirischer Vergleich zwischen Grossbritannien und Deutschland. In *Proceedings of the Conference Mensch und Computer 2005*. Kunst und Wissenschaft, Grenzilberschreitungen der interaktiven, Art, 93–101.

HOFSTEDE, G. 1986. Cultural differences in teaching and learning. *Int. J. Intercult. Relat. 10*, 301–320.

HOFSTEDE, G. 1997. *Cultures and Organizations: Software of the Mind*. McGraw-Hill.

HOFSTEDE, G. 2001. *Culture's Consequences: Comparing Values, Behaviors and Organizations across Nations* 2nd Ed. Sage Publications.

HOOK, K. 1997. Evaluating the utility and usability of an adaptive hypermedia system. In *Proceedings of the International Conference on Intelligent User Interfaces*.

INTERNATIONAL STANDARD ORGANIZATION. 1997. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs): Guidance on usability.

ITO, M. AND NAKAKOJI, K. 1996. Impact of culture on user interface design. In *International User Interfaces*, Wiley, 105–126.

KAMENTZ, E. 2006. Adaptivität von Hypermedialen Lernsystemen: Ein Vorgehensmodell fur die Konzeption einer Benutzermodellierungskomponente unter Berucksichtigung Kulturbedingter Benutzereigenschaften. Ph.D. thesis, University of Hildesheim.

KAMENTZ, E., WOMSER-HACKER, C., G. SZWILLUS, AND ZIEGLER, J. 2003. Lerntheorie und Kultur: eine Voruntersuchung für die Entwicklung von Lernsystemen für internationale Zielgruppen. In *Proceedings of the Conference "Mensch und Computer"*. 349–358.

KERSTEN, G., KERSTEN, M., AND RAKOWSKI, W. 2002. Software and culture: Beyond the internationalization of the interface. *J. Glob. Inform. Manage. 10.*

KONDRATOVA, I. AND GOLDFARB, I. 2006. Cultural interface design: Global colors study. In *Proceedings of the Move to Meaningful Internet Systems, OTM'06 Workshops*. Springer, 926–934.

KROEBER, A. AND KLUCKHOHN, C. 1952. *Culture: A Critical Review of Concepts and Definitions.* Vintage Books.

LAVIE, T. AND TRACTINSKY, N. 2004. Assessing dimensions of perceived visual aesthetics of Web sites. *Int. J. Hum.-Comput. Stud. 60*, 3, 269–298.

LINDGAARD, G. 2007. Aesthetics, Visual Appeal, Usability and User Satisfaction: What Do the User's Eyes Tell the User's Brain? *Austral. J. Emerg. Techn. Soc.* 5, 1, 1–14.

LINDGAARD, G. AND DUDEK, C. 2003. What is this evasive beast we call user satisfaction? *Interact. Comput. 15*, 3, 429–452.

LINDGAARD, G., FERNANDES, G., DUDEK, C., AND BROWN, J. 2006. Attention web designers: You have 50 milliseconds to make a good first impression) *Behav. Inform. Techn. 25*, 2, 115–126.

MARCUS, A. 2000. International and intercultural user interfaces. In *User Interfaces for All*, Lawrence Erlbaum Associates, New York, 47–63.

MARCUS, A. AND BAUMGARTNER, V.-J. 2004. A practical set of culture dimensions for global user interface development. In *Proceedings of the International Conference on Computer Human Interaction.* 252–261.

MARCUS, A. AND GOULD, E. W. 2000. Crosscurrents: cultural dimensions and global Web user-interface design. *ACM Interact. 7*, 4, 32–46.

MARCUS, A. AND GOULD, E. W. 2001. Cultural dimensions and global web design: What? So what? Now what? In *Proceedings of the Conference on Human Factors and the Web*.

MCSWEENEY, B. 2002. Hofstede's model of national cultural differences and their consequences. *Hum. Relat. 55*, 1, 89–118.

MEHTA, B. AND NEJDL, W. 2007. Intelligent distributed user modeling: from semantics to learning. In *Proceedings of the International Workshop on Ubiquitous and Decentralized User Modeling*.

NISBETT, R. 2003. *The Geography of Thought*. Free Press, New York.

NORMAN, D. A. 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, New York.

NUNALLY, J. AND BERNSTEIN, I. 1994. *Psychometric Theory. McGraw-Hill Humanities/Social Sciences/Languages.*

O'NEILL-BROWN, P. 1997. Setting the stage for the culturally adaptive agent. In *Proceedings of the AAAI Fall Symposium on Socially Intelligent Agents.*

OSHLYANSKY, L., CAIRNS, P., AND THIMBLEBY, H. 2007. Validating the unified theory of acceptance and use of technology (UTAUT) tool cross-culturally. In *Proceedings of the 21st British HCI Group Annual Conference.*

REINECKE, K. 2005. A learning platform for the education of agricultural advisers in Rwanda. M.S. thesis, Computer Science Department, University of Koblenz, Germany.

REINECKE, K. AND BERNSTEIN, A. 2007. Culturally adaptive software: Moving beyond internationalization. In *Proceedings of the International Conference on Human-Computer Interaction.*

REINECKE, K. AND BERNSTEIN, A. 2008. Predicting user interface preferences of culturally ambiguous users. In *Proceedings of the International Conference on Human Factors in Computing Systems (CHI).* ACM Digital Library.

REINECKE, K. AND BERNSTEIN, A. 2009. Tell me where you've lived, and i'll tell you what you like: Adapting interfaces to cultural preferences. In *Proceedings of the International Conference on User Modeling, Adaptation, and Personalization (UMAP)*.

REINECKE, K., REIF, G., AND BERNSTEIN, A. 2007. Cultural user modeling with CUMO: An approach to overcome the personalization bootstrapping problem. In *Proceedings of the Workshop on Cultural Heritage on the Semantic Web, International Semantic Web Conference*.

REINECKE, K., SCHENKEL, S., AND BERNSTEIN, A. 2010. Modeling a user's culture. In *Handbook of Research on Culturally Aware Information Technology: Perspectives And Models*. IGI Global.

ROGERS, P. AND TAN, J. 2008. Fifty years of intercultural study: A continuum of perspectives for research and teaching. Tech. rep., Ross School of Business.

ROSE, K 2005. Aspekte der interkulturellen Systemgestaltunllj. In *Proceedings of the Conference Mensch und Computer'05*. Kunst und Wissenschaft, GrenzUberschreitungen der interaktiven Art.

SCHMID-ISLER, S. 2000. *The Language of Digital Genres—A Semiotic Investigation of Style and Iconology on the World Wide Web*. System Sciences.

SEARS, A. AND SHNEIDERMAN, B. 1994. Split menus: Effectively using selection frequency to organize menus. *ACM Trans. Comput.-Hum. Interact. 1*, 1.

SHNEIDERMAN, B. 2002. Promoting universal usability with multi-layer interface design. In *Proceedings of the ACM Conference on Universal Usability.*

SIALA, H., O'KEEFE, R., AND HONE, K 2004. The Impact of religious affiliation on trust in the context of electronic commerce. *Interact. Comput.*

TRACTINSKY, N. 1997. Aesthetics and apparent usability: Empirically Assessing cultural and methodological issues. In *Proceedings of the International Conference on Human factors in Computing Systems (CHI'97)*.

TRACTINSKY, N., COKHAVI, A., KIRSCHENBAUM, M., AND SHARFI, T. 2006. Evaluating the consistency of immediate aesthetic perceptions of web pages. *Int. J. Hum.-Comput. Stud. 64*.

TRACTINSKY, N., SHOVAL-KATZ, A., AND IKAR, D, 2000. What is beautiful is usable. *Interact. Comput. 13*, 2, 127–145.

VENKATESH, V. AND ZHANG, X. 2010. Culture and technology adoption: Theory and empirical test of the unified theory of acceptance and use of technology (UTAUT) in the U.S. vs. China. *J. Glob. Inform. Techn. Manage. 13*, 1, 5–27.

VENKATESH, V., MORRIS, M., DAVIS, G., AND DAVIS, F. 2003. User acceptance of information technology: Toward a unified view. *MIS Quart. 27*, 3, 425–478.

VOEHRINGER-KUHNT, T. 2002. *Kulturelle Einflusse auf die Gestaltung von Mensch-Maschine Systemen*. GRIN Verlag.

YEO, A. 1996. World-Wide CHI: Cultural user interfaces, a silver lining in cultural diversity. *SIGCHI Bull. 28*, 3, 7.

ZAHED, F., VAN PELT, W., AND SONG, J, 2001. Crosscurrents: Cultural dimensions and global web user interface design. *IEEE Profess. Comm. 44*, 2, 83–103.

ZHAO, W., MASSEY, B., MURPHY, J., AND LUI, F. 2003. Cultural dimensions of web site design and content. *Prometheus 21*, 75–84.

Received April 2010; revised September 2010, October 2010; accepted January 2011