

Mengze Zhou. Impact of Cultural Markers on Localized Website Usability: A Case Study on Chinese and American Version of Multilingual Websites of MSN and Air China. A Master's Paper for the M.S. in I.S. degree. April, 2014. 58 pages. Advisor: Bradley M. Hemminger

This study investigates how the web usability and design preferences differ under different cultural contexts by examining two multilingual websites in China and the United States, i.e., Air China and MSN. The study was performed in three phases: a usability test was carried out to evaluate the performance of two localized versions upon the same user's request, and then post-test questionnaires was given out to identify user's perceptions on five main types of cultural markers, and lastly a semi-structured interview was conducted to explore user's opinion about the impact of cultural dimensions on the web design and usability. The findings indicate that cultural factors play a significant role in the way users approach their interaction with multilingual websites. And the comparison and analysis procedure can also be adapted for other websites under other cultural contexts.

Headings:

Cross-culture Usability

Cultural Markers

Cultural Dimensions

Multilingual Web Design

Website Localization

IMPACT OF CULTURAL MARKERS ON LOCALIZED WEBSITE USABILITY
A CASE STUDY ON CHINESE AND AMERICAN VERSION OF MULTILINGUAL
WEBSITES OF MSN AND AIR CHINA

by
Mengze Zhou

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in Information Science.

Chapel Hill, North Carolina

April, 2014

Approved by

Bradley M. Hemminger

Table of Contents

1. Introduction	3
1.1 Background	3
1.2 Purpose of the Study	4
1.3 Problem Statement	5
2. Literature Review	6
2.1 Current Practices in Delivering Multilingual Websites	6
2.2 Hofstede's Dimensions of Culture	8
2.3 Cultural Markers	10
2.4 Usability and Culturability	11
3. Methodology & Analytic Techniques	13
3.1 Study Set Up	13
3.2 Study Participants	14
3.3 Study Procedure	15
3.3.1 Study phase1 (p1)	15
3.3.2 Study phase2 (p2)	16
3.3.3 Study phase 3 (p3)	17
4. Results	18
4.1 Summary of Data	18
4.2 Demographics	18
4.3 Usability Test Analysis	20
4.3.1 Time on Task	20
4.3.2 Number of Clicks	20
4.3.3 User Sentiments	22
4.3.4 Post-test Questionnaire (for p1)	26
4.4 Cultural Markers Analysis	28
4.5 Analysis	31

5. Discussion.....	35
5.1 Study Implications	35
5.2 Limitations	36
5.2.1 Weakness of data.....	36
5.2.2 Weakness of analysis	36
5.2.3 Sample size and homogeneity	36
5.3 Future Work	37
6. Conclusion.....	38
Bibliography	39
Appendices.....	41
Appendix I: Informed Consent Form.....	41
Appendix II: Email Announcement	44
Appendix III: Moderate Guide.....	45
Appendix IV: Pre-test questionnaire.....	50
Appendix V: Post-test questionnaire (for P1)	51
Appendix VI: Post-test questionnaire (for P2).....	53
Appendix VII: Semi-structured interview (for P3)	55

1. Introduction

1.1 Background

With the tremendous growth of Internet, particularly the World Wide Web, as well as the global online access, multi-national corporations with ambitions on expanding market are all seeking to use multilingual websites to promote international brands, and establish a reliable, professional image to a worldwide audience. However, considering that different cultural groups can have different expectations of an easily accessed and understood website design, it is never simple to launch a culturally-competent multilingual website that meets the needs of users from diverse cultural backgrounds. To visualize this vividly in Fig. 1,

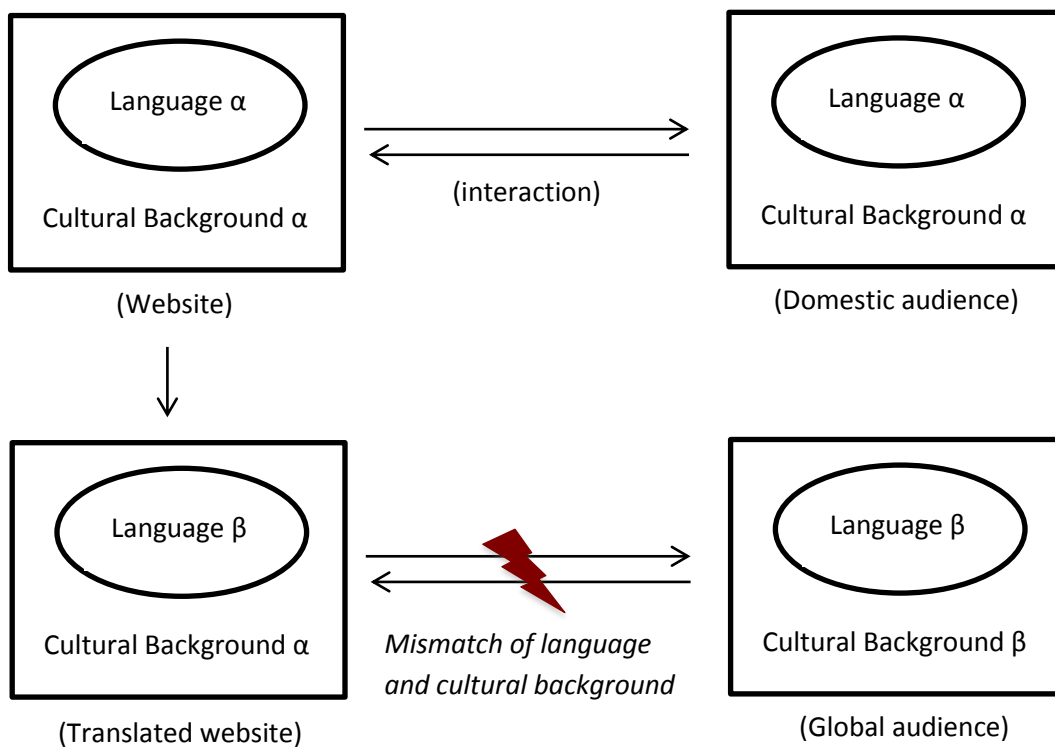


Fig. 1: Translated only website will cause a mismatch of language and cultural background

Therefore an increasing number of web designers have realized that to launch localized websites is much more than merely translating verbal components, but instead, the notion of cultural impact on web interface and its usability, has been gaining broader concern and attention, which can be beneficial to generating web-based materials, such as layout, content and tools, that are targeted toward a particular culture. In another word, an original different-looking website drawing on cultural norms should be created, of which the translated text will form only one part of the integrated whole. To gain such knowledge, designers need to identify and analyze the user needs and preferences of different cultures, the embodiment of key cultural factors on shaping differing web interface elements, i.e., cultural markers in terms of design, and then explore how the usability and the transfers of information are impacted.

1.2 Purpose of the Study

One purpose of this study is to reveal why it is of great significance to take cultural factors into consideration in web design, beginning with some reviews of the literature on diverse culture expectations and different dimensions of culture. And then by evaluating current practices of multilingual websites, specifically comparing two localized versions of the same website for China and America, this study examines different types of cultural impact on interface elements and generalizes the key culture markers that can serve as important criteria for making up localization strategies. Based on that, the study takes one step further to examine the relationship between those typical cultural markers and website usability relying on the data collected from several usability tests conducted. All the analyses above, in turn, conclude in one primary goal, which is to help designers create more effective multilingual websites for international audiences.

1.3 Problem Statement

Two case examples involved in this study are MSN and Air China, with the former being an American social media service entering into Chinese market and the latter being a Chinese airline company expanding its business in the United States. Both of them have apparently distinct looking localized websites when compared with each other. It seems that web designers have taken culture diversity into account, but what are the major underlying cultural factors that would pose this difference? Or are the two localized versions really going on the right path and rendering higher usability to its respectively targeted user group? With the overall purpose of this study as a big picture to guide, the study problems are narrowed down to:

- *How the web usability differs between the two localized versions of the same parent multilingual site, i.e., Air China and MSN, when handling the same user's request?*
- *Which cultural markers are considered as significant embodiment and how do they influence the usability of each localized version of the websites?*

2. Literature Review

2.1 Current Practices in Delivering Multilingual Websites

Although there exist difficulties and yet no recognized standard of incorporating cultural context into web design, still a considerable amount of websites are carried out with more than one language version, despite their varying levels of usability. When looking at those practices, three typical ways of handling the multilingual versions can be identified: single home sites, multi-home sites and separate sites.

The simplest approach is to deliver a single home site with partial content or sections translated into another language(s). This is often adopted by the web designers as an initial low effort solution. Users can directly see more than one language on the site page, with the translated sections appearing as links of different language to guide its target audience. Therefore, this has great limitations on both the amount and the quality of content conveyed to worldwide groups. Lately, a small number of websites begin to utilize third party services, such as Google Translate, to automatically translate blocks or even entire pages of text (see Fig. 2). Admittedly it is regarded as a big improvement by expanding the scope of text translated; however, the result is still a rough and rudimentary translation of the text, with no attention paying to the cultural impact.

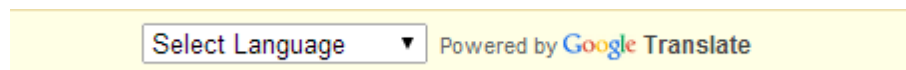


Fig. 2: Google Translate can automatically translate even entire webpages

Another alternative is to design a multi-home website that is located under one single domain name yet usually with a landing page presenting a choice of different languages (see Fig. 3). Compared with the first approach, this can well avoid the situation when users are, at the outset, exposed to the web content in an unfamiliar language, thus greatly reducing the feeling of cultural gap. But unfortunately, only a few home sites have their sub-sites delivered in different layout and design, while leaving the majority of the rest's almost the same or identical. Here what is still being ignored is the need to cater to the target cultural background.

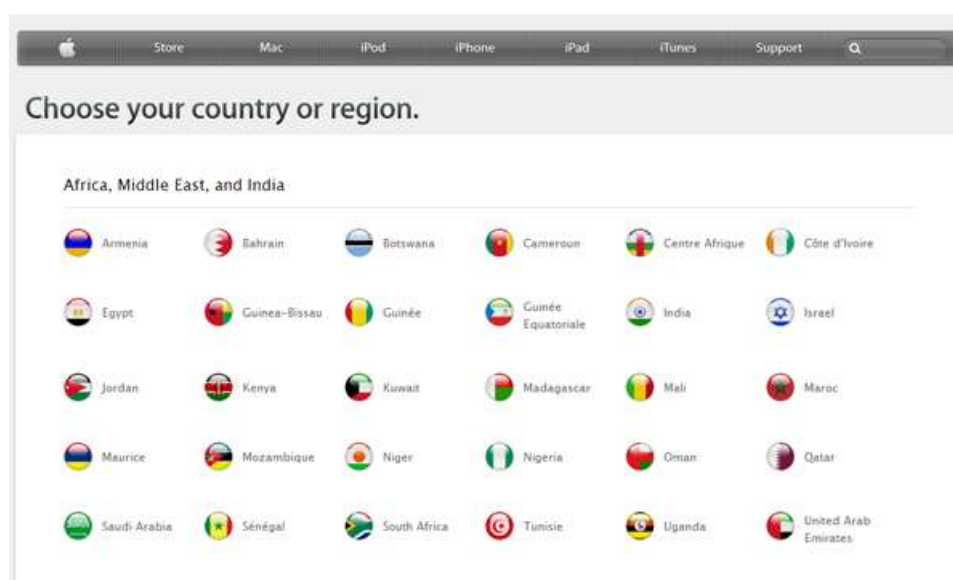


Fig. 3: A sample landing page presenting a choice of different languages

When differing language versions of a site are separately delivered, they are usually named with their own domain or sub-domain names, along with featuring different levels of variation in the web design, most of which represent the manner of taking cultural factors into account or an attempt at localization, despite the similarities among them, e.g., color schemes. It is the practice that nowadays an increasing number of businesses, especially large multinational corporations, are turning to, and this study will be focusing on the two website cases within utilizing this approach.

2.2 Hofstede's Dimensions of Culture

Hofstede describes culture as the “collective programming of the mind which distinguishes the members of one group or category of people from another”. From the viewpoint of web design, this concept can be identified as reflections of emotions, behaviors, and the way of thinking of the individuals considered as users. Furthermore, Hofstede’s approach, which contains five different dimensions of culture, appears to play an effective role in the culturalization process of global websites. Five main dimensions that manifest culturally distinctions are: Power distance (PD), Collectivism vs. Individualism (IDV), Femininity vs. Masculinity (MAS), Uncertainty Avoidance (UAI), and Long and Short Term Orientation (LTO).

Based upon this paper’s research purpose, which is to explore what differences between *Chinese* and *American* culture that have led the variation in the corresponding localized sites, first, how Hofstede compares between these two countries in terms of the five cultural dimensions is introduced. The bar chart (Fig. 4) below is taken from his website (<http://www.geert-hofstede.com>):

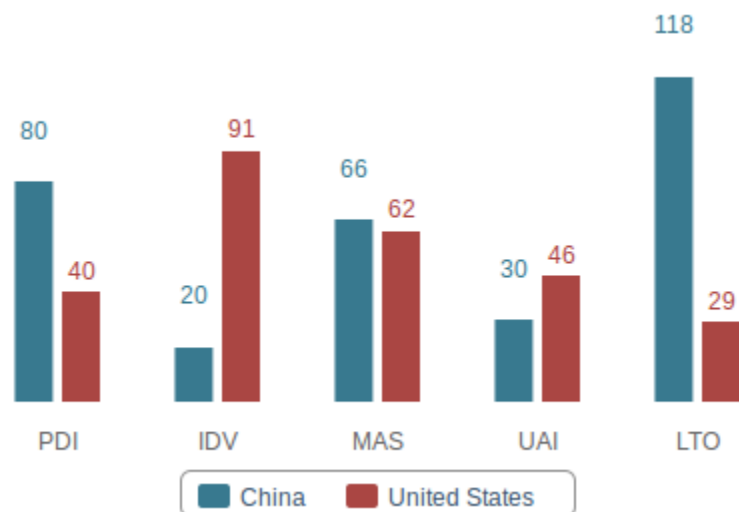


Fig. 4: Comparison on Hofstede’s five cultural dimensions between China and US

According to Hofstede, power distance (PD) refers to “the extent to which less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally”. In other word, it measures how much a culture has respect for authority. Because The PD score is inversely correlated with the degree to which subordinates participate in the decision making process, the Chinese-dominated societies, in Hofstede’s study, scored higher on power distance (with PD scores = 80) than the United States (PD score = 40). Therefore, the Chinese versions of American corporate websites should reflect higher power distance than its original English versions.

Individualism (IDV) describes the degree of interdependence a society maintains among its members. It has to do with whether people’s self-image is defined in terms of “I” or “We”. In his study, compared with China (IDV score = 20), Hofstede gave a score of up to 91 to the US, considering it a highly individual culture where people act in the interests of themselves and not necessarily of the group, therefore we should expect a more free-style web design and allow more space in organizing web content.

Masculinity-Femininity (MAS) dimension is basically based upon clear discrimination of social preferences, roles and expectations between femininity and masculinity. A high score on this dimension indicates that the society will be driven by competition, achievement and success, with success being defined by the best in field. Chinese culture has a slightly higher MAS score than the US, thus focusing a little bit more on traditional gender or age distinctions, work tasks and roles, narrowly navigation oriented to exploration while with less attention on visual aesthetics and appeals.

Uncertainty Avoidance (UAI) reflects the extent to which the members of a culture feel threatened by ambiguous or unknown situations and have created beliefs and institutions that try to avoid these. Since China is a developing country, also just like the Chinese language full of

ambiguous meanings that can be difficult for the Westerns to follow, China scores lower on UAI than the United States, which can lead to an flexible and adaptable web design with much more complexity of content and multiple types of interface controls.

For Long-Term Orientation (LTO), it measures the extent to which a society shows a pragmatic future-oriented perspective rather than a conventional historical short-term point of view. Hofstede, in his study, emphasized that eastern countries are oriented to practice and the search for virtuous behavior while Western countries are oriented to belief and the search for truth. As a result of its greatly higher LTO score, Chinese websites do not desire as much as the American websites do for immediate results, and so call for more patience to achieve navigational and functional goals.

2.3 Cultural Markers

The concept “cultural markers” was first proposed by Barber and Badre [2] to refer to “interface design elements and features that are prevalent, and possibly preferred, within a particular cultural group.” Specific cultural markers signify a cultural affiliation and denote a conventionalized use of the feature in the website. They are discovered and identified in Barber and Badre’s study that used a two stage process. The first stage involved categorizing hundreds of web sites by country, genre, and language. For the second stage, a detailed inspection was performed on interface design elements of the collected websites, and then they were able to generalize a list of recurrent design preferences by clustering websites based on their nation, language, genres, etc. Some of the typical cultural markers could be color preference, spatial layout, navigational patterns and so on.

Besides, it is also hypothesized by Barber and Badre that the elicitation of cultural markers into web design will improve the usability of the site for individuals from the culture the

website aimed at, or put it in another way, websites that contain the cultural markers of their target audience are considered more acceptable by users of their underlying culture.

2.4 Usability and Culturability

The International Organization for Standardization (ISO) defines Usability as “the extent to which the product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” It is the discipline that investigates the impact of qualitative factors in the design of web interfaces, with several consolidated guidelines and rules to evaluate whether a particular website or web-based application is usable, useful or not. Actually, usability is largely a matter of conventions, for there is no one-fits-all evaluation measure or metric. According to the definition, usability should be discussed and compared in the specified context of use. Therefore, what is usable that is perceived by users has evolved over time, and also can vary among different cultural groups. If usability is bound to culture, a question may arise of how can we transcending the boundaries to adapt to the globalized context, and to users with so different cultural backgrounds.

Therefore, according to M. Ito and K. Nakakoji, global audience interacting with a web page can be considered as a communicative action, with each phase of this process being influenced by cultural factors [38]. The two major phases are [38]:

- *Listening mode (From the perspective of user).* Toward the information a website presents, the user forms a semantic association, which subsequently leads to some level of comprehension of the information presented and;
- *Speaking mode (From the perspective of website).* Toward the intention a user forms, the website checks the applicability or realizability of the proposed action, such as content return or error, and then performs with the corresponding response.

If any of the two phases is disturbed due to the cultural gap, e.g., users misreading the information displayed, or websites unprepared for information needs of certain user groups, etc., the interaction with a multitude of international audiences is compromised, thus hampering the transfer of business messages and the delivery of products.

To this end, the notion of “Culturability” has been proposed also by Barber and Badre, which extends traditional usability to take into consideration usage preferences deriving from multiple cultural conventions, much like a new dimension in the usability matrix. It is defined as the usability in presence of influential cultural factors, i.e., cultural markers. When applied to web design, culturability is capable of capturing the cultural nuances of a targeted audience to enhance usability.

3. Methodology & Analytic Techniques

3.1 Study Set Up

The study was conducted in March 2014 on the campus of University of North Carolina at Chapel Hill. It was constructed as a within subjects design and consisted of four separate tasks for participants using either the English version or Chinese version of the two website cases, i.e. Air China and MSN (see Table 1), in a laboratory environment.

	Air China	MSN
<i>English version</i>	http://www.airchina.us/en/index.html	http://www.msn.com/
<i>Chinese version</i>	http://www.airchina.com.cn/	http://cn.msn.com/

Table 1: English and Chinese version of the two website cases, Air China and MSN

To be specific, randomly half of the participants were assigned with the Chinese version of Air China along with the English version of MSN, while the other half were told to complete tasks on the opposite version of the two sites, thus forming as control groups. The rationales for doing so are as follows:

- If no control groups were set, in other word, each participant were to take the same task consecutively on two localized versions of the same parent site, their preconceived impression from the first would inevitably cause bias and even inertial thinking on the information behavior to the second, although the design elements, such as page layout, navigational function, etc., are quite differed.
- If each participant was assigned with the same localized version of both two sites, the results collected would be insufficient, for participants loses the opportunity to

compare and express their ideas on how different cultural factors would shape different design of websites.

All participants received the same background briefing on the difference between American and Chinese culture from the perspective of Hofstede's five cultural dimension and the prevalent cultural markers in web design. And also, a multi-method approach was adopted to examine the performances of different versions of a website upon the same task by applying quantitative evaluation metrics e.g., time on task, success rate, as well as qualitative feedback from the users regarding user sentiments and satisfaction during their interaction with the two sites. Details are further explained in the following parts.

3.2 Study Participants

UNC graduate students are chosen as the target population for this study, because they are expected to have a certain level of computer operating skills that are needed for doing the usability test.

Participants were recruited by emails (see Appendix B) sent to several UNC mailing listservs, and were selected based on the eligibility criteria that the study was looking for participants who 1) were at least eighteen years old, 2) had a basic knowledge of computer skills, 3) preferably were fluent in reading both English and Chinese 4) preferably had basic knowledge on both American and Chinese culture.

There were fourteen potential participants responding to the investigator by email, and finally ten participants were selected by their satisfaction of the eligibility criteria and by the time they responded to the investigator. They were from different departments of the university, although most of them came from SILS.

All participants were assigned with a random ID number, and were randomly divided as pairs, then into two control groups (Table 2).

	Air China	MSN
Control Group A	Chinese version	English version
Control Group B	English version	Chinese version

Table 2: Two control groups are assigned with different versions of websites

3.3 Study Procedure

This study was divided into three main phases that each participant would go through, which took approximately sixty minutes:

- Four usability testing tasks given out to evaluate the performance of two localized versions of the same parent site handling the same user's request, with both performance metrics and user-reported metrics applied. (p1);
- Separate post-test questionnaires given out to examine user's perceptions on five main types of cultural markers that have embodiment in the design of both two websites. (p2);
- A semi-structured interview conducted to explore user's opinion about the impact of cultural dimensions on the web design and interface usability. (p3).

3.3.1 Study phase1 (p1)

Experimental approach, specifically running basic usability testing, is adopted for the study phase one. Ten participants are divided into two control groups with each respectively examining one localized version of the two case examples. In each group, one will be asked to complete four tasks, first two on Chinese version of either MSN or Air China website and the rest two on English version of the other website

Task 1: Checking the flight schedule information. (Air China)

Task 2: Checking the airplane model information. (Air China)

Task 3: Checking the score of one NBA game (MSN)

Task 4: Checking today's gold price (MSN)

Rylstim Screen Recorder 1.5 is used to record the screen capture video for the entirety of each session. Following each session, the duration time of each task and the number of mouse clicks are calculated. And after completing the four tasks above, each participant will be asked to fill in a post-test questionnaire (for p1) regarding their subjective impressions on the websites.

3.3.2 Study phase2 (p2)

Questionnaires are adopted for study phase 2 to collect quantitative data. To reduce the complexity of the implementation, only five major categories of cultural markers are used, which are: language/text, visual/graphic elements, colors, page layout and navigational function. Details are illustrated in the table 3 below. After each participant has completed those four tasks, they will be given the separate questionnaires to evaluate or express their impression and opinion on cultural markers of the two websites they've browsed, e.g., is this cultural marker easily noticeable and do some of them make the website easier to understand? In this process, participants are asked to revisit the previous websites and also encouraged to think aloud when filling in the questionnaire.

Language/Text	Visual/graphic Elements	Colors	Page Layout	Navigational Function
Text density	Logo/brand	Color intensity	Page length	Menu-bar intensity
Text size	Image density	Color relations	Centered or not	Button intensity
Text style integrity	Image size	Color diversity	Regular or not	Link intensity
Language integrity	Image resolution			Flexibility
Content integrity	Video display			
	Animation			

Table 3: Five major categories of cultural markers are examined in this study

A scale of 1 to 5 is also used in rating the extent of user's perception on each of the five cultural markers: 1 = not perceptible 2 = hardly perceptible, 3 = perceptible to some extent, 4 = clearly perceptible and 5 = strongly perceptible.

3.3.3 Study phase 3 (p3)

Relying on the literature frameworks that have been proposed by Hofstede, the cultural dimensions that are used for verification are: Power distance (PD), Collectivism vs. Individualism (IDV), Femininity vs. Masculinity (MAS), Uncertainty Avoidance (UA), and Long and Short Term Orientation (LTO)

After each participant has submitted the questionnaires, they will be interviewed about their perspective on the impact of cultural dimensions on the web design and interface usability. The five cultural dimensions along with its explanation and effects on web design are explained, and will also be printed out and handed out. The participants will be asked to talk about how much importance s/he would consider for each of the five cultural dimensions, i.e., to what degree every cultural dimension has its embodiment on the two case examples. The responses to all those questions should reflect how the participant sees the significance and the influence of the five cultural dimensions.

4. Results

4.1 Summary of Data

All the main data collected through the study include: demographic information from pre-test questionnaire, performance metrics recorded during usability test, user-reported metrics from post-task interviews and post-test questionnaire I (for p1), user's rating regarding cultural markers perceived in web interface from a separate post-test questionnaire II (for p2), and user's reflection on the impact of Hofstede's cultural dimensions on web design from a semi-structured interview. Table 4 shows the captured data and their corresponding data types.

Data Resource	Data Analysis	Data Type	Variable Property
Pre-test questionnaire	Summary of demographic info	Qualitative	Mixed
Usability test	Statistical analysis of performance metrics	Quantitative	Mixed
Post-test questionnaire I (for p1)	Statistical analysis of user-reported metrics	Quantitative	Continuous
Post-test questionnaire II (for p2)	Statistical analysis of user's rating on cultural markers	Quantitative	Continuous
Semi-structured interview	Summary of the impact of cultural dimensions on web design	Qualitative	

Table 4: Data that are captured in this study and their corresponding data types

4.2 Demographics

All demographic information is obtained from the pre-test questionnaire given out to participants prior to the usability test. Based on the analysis of collected data, the gender distribution is 40% male and 60% female. As expected, the female students show a bit more

interest on this study considering the topic is culture-related. The age of all ten participants recruited ranges from 23 to 29 and they are all full-time UNC graduate students, half of which have previously participated in a usability test. Also, the result shows that all participants can read and understand both written English and Chinese, with three of them stating that they are familiar with both American and Chinese culture, one more familiar with the American and the rest six knowing more about Chinese culture. Overall this is somewhat uneven in terms of the cultural groups that participants belong to, which will may cause some bias in information behaviors and user feedback on the two localized versions of websites. Although the best experimental condition is that all participants are equally familiar with both cultures, practical situation of this nature is to some extent remedied via pre-test background briefing on typical differences between American and Chinese culture in view of Hofstede's five cultural dimensions as well as the major cultural markers in web design.

Because the two multilingual website cases involved in this study are: 1) official website of airline and 2) information web portal, some domain-related demographic questions are also raised. According to the responses, all participants are experienced with both two types of websites. Two fifths of participants frequently refer to the website of airlines for services like online flight booking, flight schedule checking, etc., and seven tenths say that they browse web portals for information very often. Besides, all participants have been to any multilingual websites before and more than half state that they are quite used to and would frequently visit such websites. Those results indicate there's a good chance that the performance of websites handling user's request is dependent on the usability on website's own side, with little impact due to the inexperience on participant's side. Therefore, all data collected in the pre-test questionnaire well support that the participants recruited in this study can serve as a good

subject pool which lays a solid start for the usability test, thus also enhancing the accuracy and credibility for the study findings.

4.3 Usability Test Analysis

Each applied evaluation measures, ranging from performance metrics to user-reported metrics, in reference to the four tasks overall will be discussed, and then in this section will present a breakdown of each task with the main issues highlighted.

4.3.1 Time on Task

After completing the test, *time on task* proved to be an ineffective metric primarily given the different page loading speed for the English and Chinese versions of both Air China and MSN. What has been ignored before test is that in order to maximize the web usability and improve user experience for target audience, the localized version of both two multilingual websites are held separately on regional servers, therefore for this study, which takes place in the US, participants will have to expect more responding time for connecting to the Chinese version. This extraneous noise on interfering the accurate time on “task” is so strong that has to be removed. In addition, time on task when combined with participant’s think-aloud during the test would not reflect real world usage, and for the purpose of this study, data gained from think-aloud would be more useful when participants made reflection on the design elements that are culturally embedded.

However, it can still be inferred that users were generally satisfied with the time it took to complete tasks, as none of them complained about task length, and found the four tasks to be reasonable in difficulty.

4.3.2 Number of Clicks

Compared with time on task, the metric *number of clicks* is more useful and effective in measuring the level of website usability, for it does not count unwanted waiting time when the

page is being prepared. Under the assumption that almost every mouse click on the screen is task-related, the fewer number of clicks, the faster and easier users would find in completing tasks, thus indicating higher usability. The Fig. 5 shows how the number of clicks distributed in the two versions of both Air China and MSN in response to the four tasks that ten participants conducted.

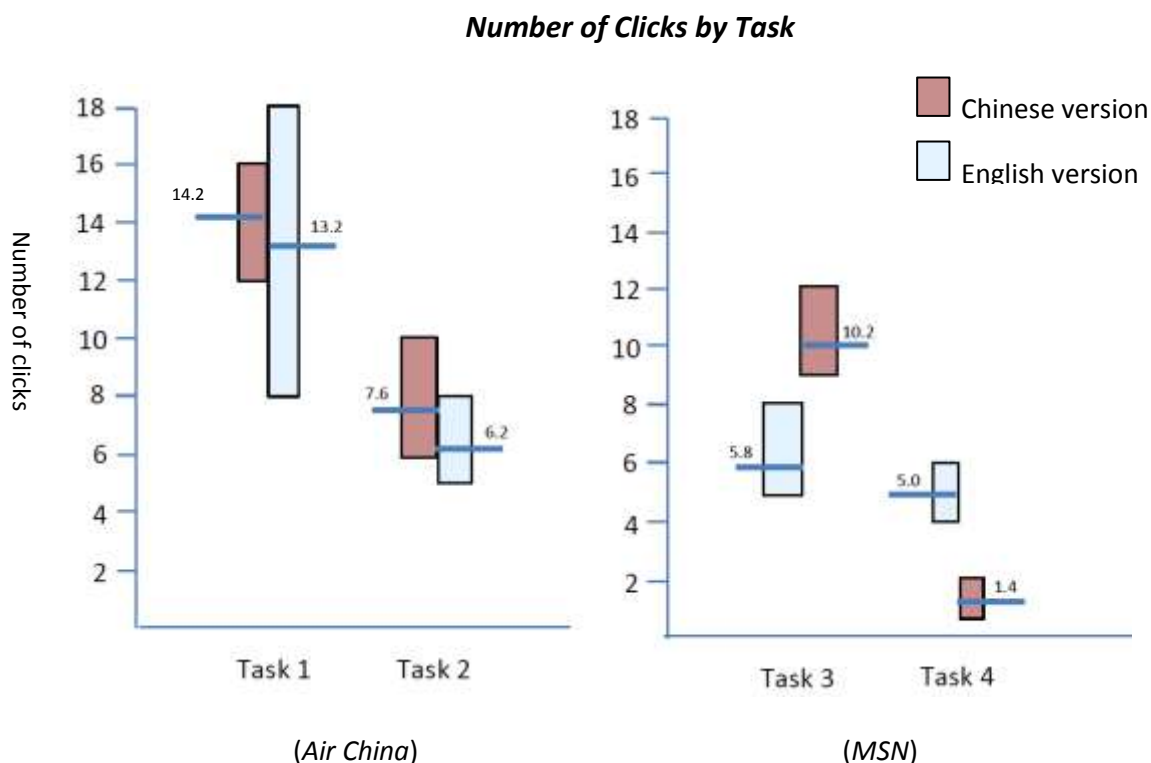


Fig. 5: Number of clicks by task distributed in the two versions of both Air China and MSN

As depicted in the chart above, the number of clicks differs in the two localized versions upon the same task. The average of number of clicks of the five participants', from control group A, completing task 1, task 2 and task 3 is all greater than that of the other control group, demonstrating that the English version of both Air China and MSN have a higher performance than their corresponding Chinese version. However, there's one outlier, task 4, which is to check today's gold price within the web portal MSN. The number of clicks when participants search on its Chinese version is abnormally lower and the reason lies in one special spot, which has

something to do with cultural factors, of the web structure. In Chinese culture, investment on gold is for a long time favored and receives much more attention than that in American culture. Therefore, the link to the section “Gold” is brought into the main navigation bar on top of its home site; while in comparison, it cannot be easily found in the English version, as it is hidden under the “Money -> Market -> Commodity -> Metal”. Lower number of clicks indicates a good practice for incorporating target cultural background into the page layout and interface design of the localized site. From this, one more interesting questions arises: what is the reason for causing more mouse clicks on the Chinese version of both multilingual websites for the first three tasks? Is it because they lack or not doing the proper way of customizing the website into the cultural background, or is such a result of any already-embedded cultural conventions? The answer is discussed and proposed later.

4.3.3 User Sentiments

User sentiments were coded much in the same way as errors, in that a tagged notes or quotes of participants that appeared as especially positive or negative concerning the website in completing tasks. Therefore it is like using quantitative method to process the qualitative data or descriptive feedback that users convey during the usability test. As a user-reported metric, user sentiment provides valuable complementary information to task performance in evaluating web usability and can serve as a good supplementary data to explain why the number of clicks varies in the two localized version of Air China and MSN. If more negative comments go along with more mouse clicks, that would indicate poor usability; alternatively, if few negative comments go along with more mouse clicks, it indicates that participants did not feel frustrated about the paths or steps made on the website, which may suggest that there’s no gap or mismatch between the localized version and its target cultural context, which is also a significant aspect that needs to be identified in this study.

Based upon tagging the collected participants' feedback, total numbers of positive and negative sentiments grouped by task are presented in the Fig. 6 below:

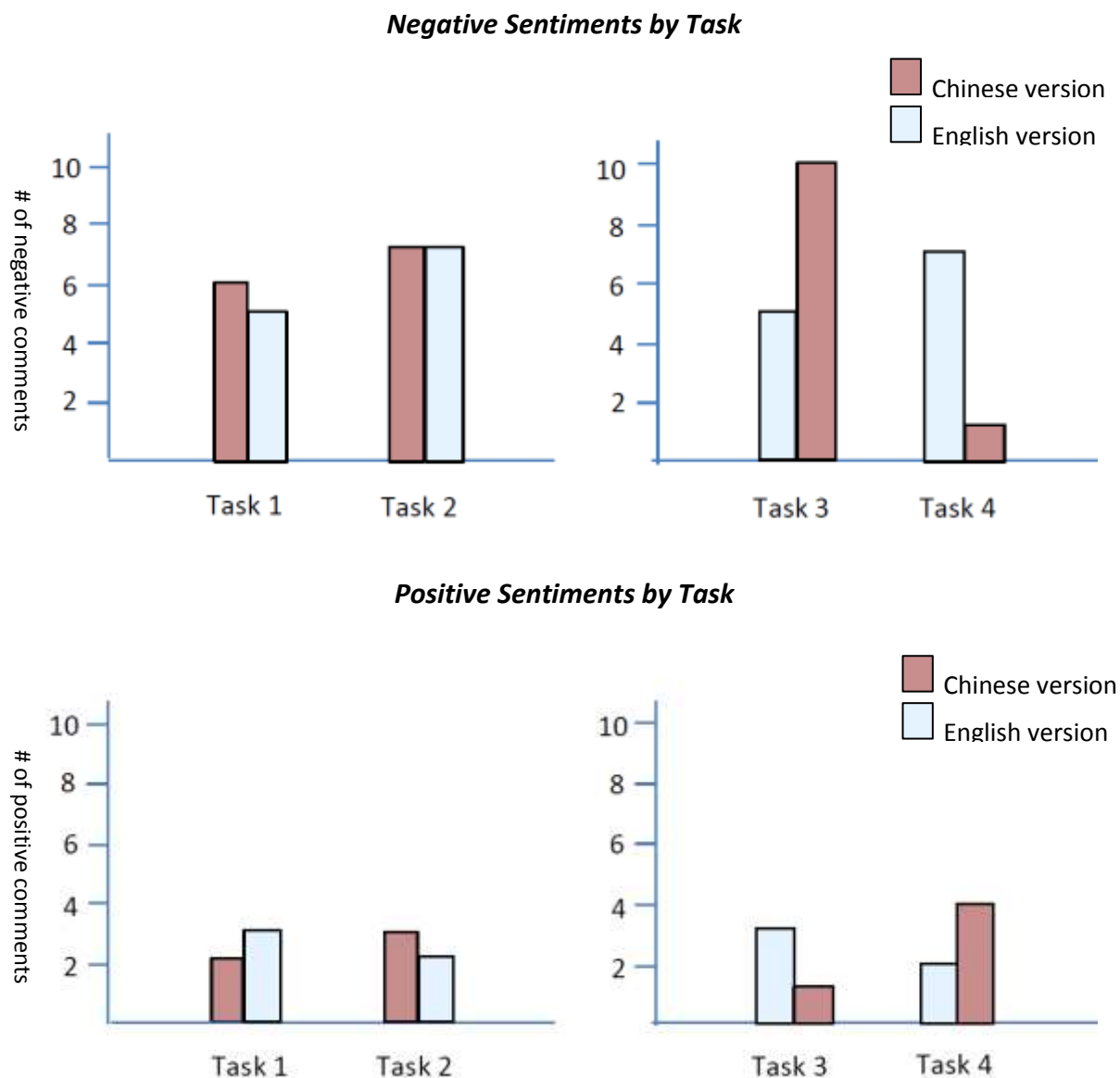


Fig. 6: User sentiments by task distributed in the two versions of Air China and MSN

While comparing two charts does show that negative comments overwhelm the positive ones, this may be due to the nature of the usability test. Usually participants come in expecting

to provide critical feedback on the website or system and are less likely to offer unsolicited positive comments.

The outcomes above of analyzing user sentiments are largely as expected. For task 1, task 2 and task 3, the average of number of clicks on the Chinese version is all greater than that on the English version, so correspondingly, more negative comments and fewer positive ones are collected during participant's interacting with the Chinese version of both two websites. If the deviation of the number of mouse clicks between the two versions is remarkable, like task 3, then the difference of the total of negative comments is also significant. And for task 4, the situation is reversed, so the number of negative comments on the English version well exceeds that on the other site.

To step further, a closer look at the top frequent and typical negative/positive user sentiments reveals more detail about what caused the problems and impacted the web usability:

The Chinese version of Air China:

- Its main navigation bar contains so many links of sub-sections that users can be easily diverted when searching for particular section. *(negative)*
- There are dead links on page "Inflight Entertainment" that block users from access and should have directed to related information, as is shown in the screenshot Fig. 7 below. *(negative)*

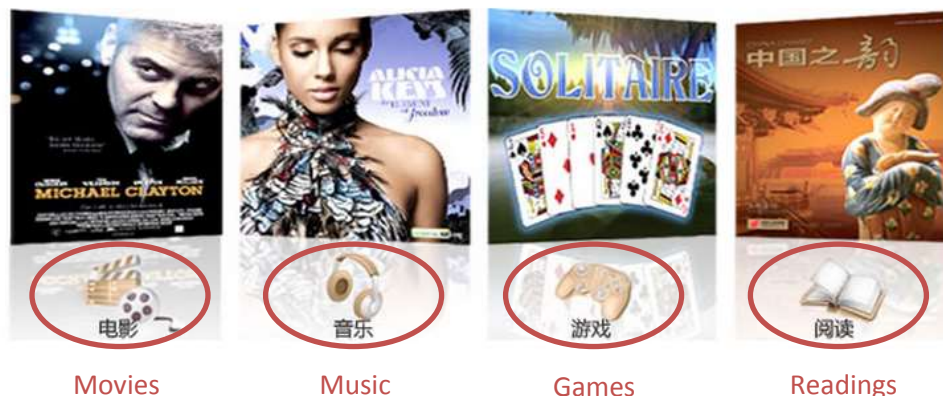


Fig. 7: Dead links on page “Inflight Entertainment” in the Chinese version of Air China

- In its flight search module, popular cities are directly suggested in the form of drop-down menu after the mouse clicks on the text field. (*positive*)

The English version of Air China:

- In its flight search result display page, it shows all the flight information for a time period of the user-specified date +/- 3 days, which is convenient, while not allowing users to browse the schedule that is beyond three days, which is inflexible. See the Fig. 8 below. (*negative*)



Fig. 8: Inflexible selection of schedule date in the English version of Air China

- The navigation bar is simple, plain and well-structured. (*positive*)

The Chinese version of MSN:

- Some hot sections lack of structured grouping, e.g., under the section “Basketball”, no further grouping are provided such as “NBA”, “CBA”, etc., which makes the information simply messy. (*negative*)

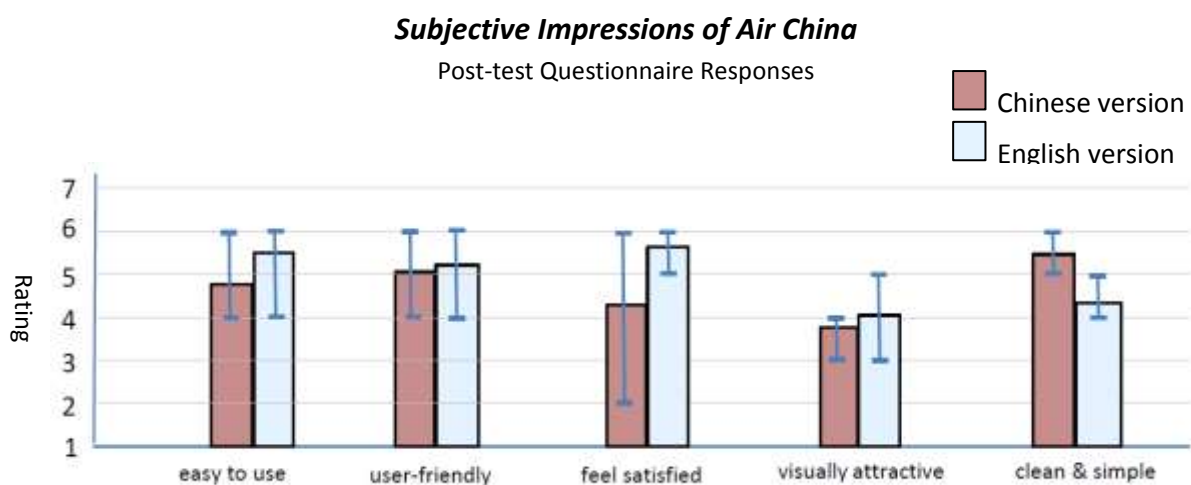
- Taken into account that “gold” is China’s top investment point, its navigational link is listed in a conspicuous place for users to easily get access to. (*positive*)

The English version of MSN:

- Its organization of web content is close-knit with its clear navigational structure.
(*positive*)

4.3.4 Post-test Questionnaire (for p1)

The post-test questionnaire (for p1) was given out to participants right after they completed the four tasks, which is to make an additional quantitative assessment on user-reported metrics based on their previous experience and impression on the two localized versions of both Air China and MSN. All participants’ feedback is measured using a 7-point Likert-type rating scale (1=Strongly Disagree, 7=Strongly Agree), and the data collected is integrated in the following charts Fig. 9.



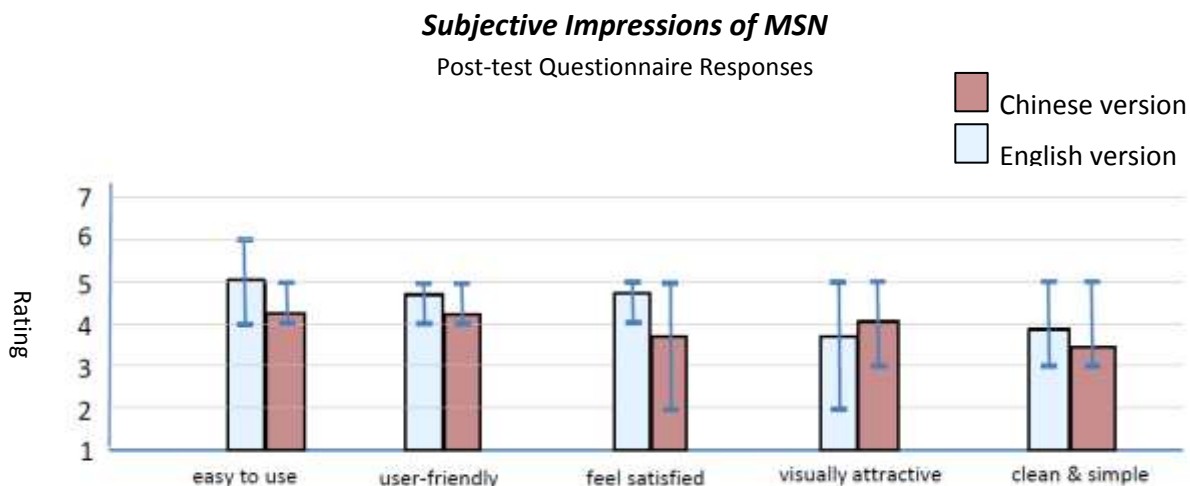


Fig. 9: Post-test questionnaire I responses

The colored bar stands for the average score for each corresponding criteria, and the I-shaped line represents the minimum and maximum score received. The results above show that participants are generally more satisfied with the English version of both Air China and MSN, which just matches with the other evaluation metrics illustrated in the prior parts, thus consolidating the conclusion that for the two multilingual websites, Air China and MSN, the English versions of both sites have higher usability than the Chinese versions do. Yet there's also one interesting point to note: as is shown in the charts, participant's scores on the first two questions regarding the ease-of-use and user-friendliness of the two versions are really close, but for the question on the satisfaction, substantial divergence appears for the Chinese version of Air China (minimum score = 2, maximum score = 6) and MSN (minimum score = 2, maximum score = 5), which drags down the average score. From the scores for the first two questions, it can be inferred that no great difficulties are encountered in experiencing both two versions of site and completing the tasks, then is there any other reason that impacted the score of general "satisfaction"? Could this concern with cultural factors?

4.4 Cultural Markers Analysis

In order to probe into the answers to those questions brought by the results of usability test, participants were asked to do a follow-up separate post-test questionnaire (for p2), which includes a checkpoint listing of five major types of cultural markers prevalent in conventional websites, to verify their perception (on a 1 to 5 scale, with 1=too few/low or not perceptible, 2=few/low, 3= moderate/right, 4=high/many 5=too high/too many or strongly perceptible) on certain design preferences under a particular cultural context. And the data collected clearly indicate that there is a variation in some cultural markers between the Chinese and English versions of the same parent site, either Air China or MSN, as is summarized below the key findings based on statistical analysis.

Language/Text: (see Table 5)

Average score	Air China		MSN	
	Chinese version	English version	Chinese version	English version
Text density	2.6	4	3.4	4
Text size (off-size?)	3.4	1.8	3.4	1.6
Text style integrity	3.4	4	4.4	3.6
Language integrity	4.4	4.2	4.4	4.4
Content integrity	4.6	3.8	3.6	3.6

Table 5: Comparison on Language/Text of both two versions of Air China and MSN

Through the comparison of the results shown in this chart, some important conclusions can be highlighted:

- For both Air China and MSN, the English version site has higher density as well as notably smaller size of text displayed on page, the reason for which is that both two localized websites have taken into account the difference between English and Chinese characters. If this factor was improperly integrated into the website for its target cultural group, say, enlarging the text size while also increasing text space in the English version, its current high usability would much likely be compromised.

- Regarding the integrity of text style, display language and web content, the Chinese version of both two multilingual websites received higher scores, indicating a more regulated and controlled web design, compared with a more free-style one embodies in the corresponding English version. This has something to do with the cultural dimension Collectivism vs. Individualism (IND), which will be further discussed in the next section.

Visual/graphic Elements: (see Table 6)

Average score	Air China		MSN	
	Chinese version	English version	Chinese version	English version
Logo/brand	4.2	4	3.8	3
Image density	4	3	4	4
Image size	3.8	2.4	2.8	3.6
Image resolution	4.4	3.2	4	4.6
Video display	1.2	1	1.2	3.2
Animation	3	2.8	2.4	3.2

Table 6: Comparison on Visual/graphic Elements of both two versions of Air China and MSN

Through the comparison of the results shown in this chart, some important conclusions can be highlighted:

- The official logo or brand for both Air China and MSN is emphasized more on their Chinese versions, which is caused by China's higher Power Distance (PD) in its culture dimension than American's.
- For other visual/graphic elements, i.e., image, video and animation, there's a disagreement according to the statistical outcome: the Chinese version of Air China features more visual elements than its English version; while for MSN, the English version has more. This may explain the interesting question raised in the previous part about why the score of user's satisfaction for the Chinese version of Air China fluctuates. One participant who is more familiar with Chinese culture once mentioned during the test that the images and animation on Air China's Chinese version seem too crowded and even overwhelming, thus causing a little bit discomfort, therefore if there's a

mismatch between the cultural marker in interface design and the cultural context where user's mental model builds on, the web usability and user experience would be affected.

Colors: (see Table 7)

Average score	Air China		MSN	
	Chinese version	English version	Chinese version	English version
Color intensity	3.6	3.6	3.6	3.4
Color relations	3.6	3.8	4	4.6
Color diversity	3.4	3.2	2.6	2.4

Table 7: Comparison on of Colors of both two versions of Air China and MSN

Through the comparison of the results shown in this chart, some important conclusions can be highlighted:

- In general, the use of color between both Chinese and English version of Air China and MSN is in close proximity. This is probably because the two versions need to apply the similar color scheme to represent and enhance the corporate identity.
- One thing to note is that the color relation for the English version of both Air China and MSN is at the same time higher than that of the Chinese version, so combined with the result that the usability of English version is also higher, it may be inferred that using colors that are adjacent to each other can be to some extent helpful in improving user experience.

Page Layout: (see table 8)

Average score	Air China		MSN	
	Chinese version	English version	Chinese version	English version
Page length	1.4	2	4.4	4.4
Centered or not	Y	Y	Y	Y
Regular or not	Y	Y	Y	Y

Table 8: Comparison on Page Layout of both two versions of Air China and MSN

Through the comparison of the results shown in this chart, some important conclusions can be highlighted:

- It shows that the page layout is quite consistent for the two localized versions of Air China and MSN: the page length is almost equivalent, and each version adopts the center-aligned fixed-width layouts, with the content displayed by means of regular “blocks of data”.

Navigational Function: (see Table 9)

Average score	Air China		MSN	
	Chinese version	English version	Chinese version	English version
Menu-bar intensity	4	3.2	3.6	3.4
Button intensity	3	3.6	3.4	3.2
Link intensity	3.4	4	4	4.4
Flexibility	3.2	4	3.2	3.4

Table 9: Comparison on Navigational Function of both two versions of Air China and MSN

Through the comparison of the results shown in this chart, some important conclusions can be highlighted:

- Regarding the intensity of menu-bar, the Chinese version of both two parent sites scores higher, meanwhile combining with its lower flexibility indicates a more rigid navigational structure. This tall hierarchy of organizing the website also implies a higher Power Distance (PD) in Chinese culture.
- For the intensity of buttons and links, however, the English version site receives a higher score, meaning that more buttons and links tend to be directly appearing on its pages, which can be explained with American’s desire for immediate results, i.e., lower Long and Short Term Orientation (LTO) in terms of cultural dimension.

4.5 Analysis

The data about the impact of Hofstede’s five cultural dimensions on web design and interface usability is collected via a semi-structured interview with participants to explore their

opinion and thoughts combined with their experience on the two multilingual websites. Below is the summary of major constructive results.

Impact of Power Distance (PD) on web design and usability:

- Higher Power Distance will focus strongly on authority, therefore the official logo or brand will be emphasized more in the design of the target-culture localized site, yet this would exert no significant differences regarding the web usability.
- Higher Power Distance will also focus on highly-structured access to web content, therefore tall hierarchy of organizing the website is expected, which may impact the speed of user's information searching. If users from the cultural background of low Power Distance are given the website with a mismatched high PD, it may take longer for them to complete the tasks.
- Higher Power Distance might put certain restrictions or blocks on some parts or sections of a website for users to freely get access to, thus to some extent, lowering the usability. E.g., on the Chinese version of Air China, participants cannot click into the "Inflight Entertainment" section for information, which can be a reflection of higher Power Distance.

Impact of Collectivism vs. Individualism (IND) on web design and usability:

- Higher Individualism will usually bring a more free-style design of website with not so strict demand on the integrity of text style, display language and web content, while on the contrary, higher Collectivism can be identified in a more traditionally designed website. And for web usability, it is not directly or remarkably impacted, yet perhaps users coming from the culture of high Collectivism may tend to be more likely to get lost in the website within a high Individualism cultural context.

Impact of Masculinity vs. femininity (MAS) on web design and usability:

- Although the MAS score is very close for the Chinese and American culture and no clear distinction on MAS between the localized version of two multilingual websites, considering that higher Masculinity focuses more on the functionality of a website while higher Femininity would emphasize more on visual aesthetics, an assumption can be made that the masculine site is likely to have superior usability to the feminine one in terms of accuracy, speed and satisfaction levels, especially male user's.

Impact of Uncertainty Avoidance (UA) on web design and usability:

- Higher Uncertainty Avoidance will focus on the prevention of user errors, such as by providing well-structured, clean and minimum navigation structure to prevent users from getting lost. Combined with the positive sentiments made on the English version of both Air China and MSN: *"The navigation bar is simple, plain and well-structured."*, *"Its organization of web content is close-knit with its clear navigational structure."*, it can be inferred that English versions is put under the cultural context that values high UA, which matches well with Hofstede's theory that American culture scores higher on UAI than Chinese culture.
- Besides, an assumption can be made that low uncertainty avoidant users with cultural background of low UA can perform better on a high uncertainty avoidant site; while for the user group of high UA, the performance can be just the opposite when they are experiencing a low uncertainty avoidant site. Combined with previous finding that the English version of both two multilingual website cases has higher usability over the Chinese version, it could suggest that the high uncertainty avoidant site is better in usability in terms of user's satisfaction levels.

Impact of Long and Short Term Orientation (LTO) on web design and usability:

- Lower LTO will focus on helping users get access to information like top searches in order to satisfy the desire for immediate results, therefore more buttons and links tend to occur more frequently on web pages, which might suggest higher usability.

5. Discussion

5.1 Study Implications

Based on the analysis of data collected through three phases of the study, the results suggest that integrating cultural factors is of great significance in shaping better usability in the multilingual web design.

There is a non-negligible correlation between Hofstede's five cultural dimensions and localized version of websites. A better web design should be aligned to or accommodate at least one or more cultural dimensions, which will result in an increase in usability of targeted users. Second, as a concrete manifestation of invisible cultural dimensions, cultural markers perform as an efficient and effective strategy to address local audiences and increase usability of multilingual websites if used appropriately. Otherwise, if cultural markers are applied unselectively, or only the characteristics of cultural minorities are represented, the localization process might have a negative impact on user's understanding and perception.

Therefore, in today's era of globalization, how to define and make an effective web design process suitable for the multi-cultural audience should be addressed in immediate need. Because traditional usability guidelines overlook differing expectations and preferences prevailing in varied cultural context, to have a better understanding and recognition on the role and influence of the cultural factors on web usability can be of great benefits to big businesses which seeks to utilize multilingual websites or web-based application to extend to international markets. And directly, based on the study results, web designers can regard this as an easy-to-check reference to see if the current or under-prototyping websites fit the characteristics of the

cultural dimensions, and include the appropriate cultural marks in specific cultural group, thus creating or redesigning more targeted localized websites.

5.2 Limitations

5.2.1 Weakness of data

There are only two case examples to be examined in this study, which lacks the comprehensiveness of the other categories of websites. And also because it is possible that those two cases are under-represented, the data results collected may be one-sided, unable to fully describe the embodiment on web design of differing cultural dimensions between China and the United States, as well as user's perception on cultural markers of the two countries and their impact on web usability.

5.2.2 Weakness of analysis

Due to the limited time frame and small scope of this research study, this study only focuses on Hofstede's five cultural dimensions, and five most typical cultural markers. Therefore, the analysis on them may not cover the full picture of the impact on web usability of cultural factors.

Besides, cross-culture usability is more than just about the cultural markers, but instead, the complex and dynamic cultural contexts in which the localized websites are targeted must also be thoroughly explored.

5.2.3 Sample size and homogeneity

The small sample size and the fact that all participants were educated tech savvy university graduate students make it likely that there may be differences between these results and those for a more diverse population, especially including older adults.

5.3 Future Work

Larger scale studies are needed to investigate systematically how cultural markers work in multilingual Web design: researching cultural markers originating in as many cultures as possible, recruiting more test subjects from different local cultures, testing more representative multilingual websites, and so on.

6. Conclusion

This study probes into the relationship between cultural factors and the multilingual web design. Two case examples, i.e., MSN and Air China are used to verify the significance of cultural dimensions, examine user's perceptions on different categories of cultural markers, and evaluate the impact of several main cultural markers on web usability. The primary research approaches to take in this study are Questionnaire, Survey, and Experimental Usability Testing. By extending traditional usability to take into consideration usage preferences deriving from multiple cultural conventions, the research results will be greatly beneficial to those web designers who seek efficient and effective strategies to create more targeted localized websites, and evaluate and redesign current practices.

Bibliography

1. Barber, W. and Badre, A. Culturability: The Merging of Culture and Usability in Proceedings of the 4th Conference on Human Factors and the Web, 1998.
2. V. Evers and D. Day. The role of culture in interface acceptance. 260–267. Chapman and Hall, Ltd., 1997.
3. Sheppard, C. and Scholtz, J. (1999). The Effects of Cultural Markers on Web Site in Proceedings of the 5th Conference on Human Factors and the Web, 1999.
4. Amant, K. S.. A prototype theory approach to international web site analysis and design. Technical Communications Quarterly, 14(1), 73-91, 2005.
5. A. Marcus and E.W. Gould. Cultural dimensions and global web user-interface design: What? so what? now what. Proceedings of the 6th Conference on Human Factors and the Web, 1–15, 2000.
6. Tsiriktsis. Does culture influence web site quality expectations?: An empirical study. Journal of Service Research, 5(2):101, 2002.
7. Smith, A., Dunckley, L., French, T., Minocha, S., & Chang, Y.: A process model for developing usable cross- cultural websites, Interacting with Computers. Interacting with Computers, Volume 16, Issue 1, 63-90, 2004.
8. Sun, Huatong. "Building a culturally-competent corporate web site: an exploratory study of cultural markers in multilingual web design." Proceedings of the 19th annual international conference on Computer documentation, 2001.

9. Hofstede, G.. Dimensionalizing cultures: the Hofstede model in context. Online Readings in Psychology and Culture, 2(1), 2011.
10. A. Smith, L. Dunckley, T. French, S. Minocha, and Y. Chang. A process model for developing usable cross-cultural websites. Interacting with Computers, 16:63–91, 2004.
11. Hillier, M.: The role of cultural context in multilingual website usability. Electronic Commerce Research and Applications, 2(1), 2-14, 2003.
12. Ford, Gabrielle, and Paula Kotzé: Designing usable interfaces with cultural dimensions. Human-computer interaction-INTERACT 2005. Springer Berlin Heidelberg, 713-726, 2005.

Appendices

Appendix I: Informed Consent Form

Consent Form Version Date: 2014-03-05

Title of Study: Impact of Cultural Markers on Localized Website Usability, A Case Study on Chinese and American version of multilingual websites of MSN and Air China

Study Conductor: Mengze Zhou

Faculty Advisor: Bradley M. Hemminger

UNC-Chapel Hill Department: School of Information and Library Science

Study Conductor Email Address: z mz1107@live.unc.edu

Faculty Advisor Email Address: b mh@ils.unc.edu

What are some general things you should know about studies?

You are being asked to take part in a study. To join the study is voluntary.

You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this study.

You will be given a copy of this consent form. You should ask the study conductor named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?

The purpose of this study is to examine the impact of key culture markers (referring to “interface design elements and features that are prevalent, and possibly preferred, within a particular cultural group”) on the usability of the two multilingual website cases: MSN and Air China, with the former being an American social media service entering into Chinese market and the latter being a Chinese airline company expanding its business in the United States. Both of them have apparently distinct looking localized websites when compared with each other.

Are there any reasons you should not be in this study?

You should not be in this study if:

You are younger than 18 years old.

How many people will take part in this study?

If you decide to be in this study, you will be one of the ten people in this study.

How long will your part in this study last?

Participation will consist of a single session that will last about 60-90 minutes.

What will happen if you take part in the study?

If you agree to participate, we will arrange a mutually agreeable time and place to conduct the study. The study will consist of a pre-test questionnaire, a series of tasks that you will be asked to perform, a brief post-test interview, and a post-test questionnaire about your experience. The study will NOT be recorded, but we will take hand-written notes on your actions and responses. This data will not be shared outside an educational setting, and will in no way contain personally identifying information.

For any reason, you may choose not to answer any question that is part of the study and you may leave at any time.

What are the possible benefits from being in this study?

You may not benefit personally from being in this research study.

What are the possible risks or discomforts involved from being in this study?

We believe the risks in this study to be no more than those encountered in everyday life. There may be uncommon or previously unknown discomforts. You should report any problems to the researcher.

How will your privacy be protected?

We will not use your name or other personally identifying information in our study. Any personal identifiers during participant recruitment will be associated with randomly assigned id number, coded with a linkage file stored separately. And during the test, the screen recording tool will only record your mouse move and clicks when conducting the tasks, no face photographic images will be recorded. The raw research data we collect will be stored on our computers and on university servers. After the analysis for this project is completed, we will delete and/or destroy the originally collected raw data.

Participants will not be identified by name in any report or publication about this study. Although every effort will be made to keep research records private, there may be times when federal or state law requires the disclosure of such records, including personal information. This is very unlikely, but if disclosure is ever required, UNC-Chapel Hill will take steps allowable by law to protect the privacy of personal information. In some cases, your information in this research study could be reviewed by representatives of the University, research sponsors, or government agencies (for example, the FDA) for purposes such as quality control or safety.

What if you want to stop before your part in the study is complete?

You can withdraw from this study at any time, without penalty. The investigators also have the right to stop your participation at any time. This could be because you have had an unexpected reaction, or have failed to follow instructions, or because the entire study has been stopped.

What if you are a UNC student?

You may choose not to be in the study or to stop being in the study before it is over at any time. This will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this study.

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this study. If you have questions about the study, complaints, concerns, or if a study-related injury occurs, you should contact the faculty advisor listed on the first page of this form.

Title of Study: Impact of Cultural Markers on Localized Website Usability, A Case Study on Chinese and American version of multilingual websites of MSN and Air China

Study Conductor: Mengze Zhou

Faculty Advisor: Bradley M. Hemminger

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I meet the qualifications for the study. I voluntarily agree to participate in this research study.

Signature of Study Participant

Date

Printed Name of Study Participant

Signature of Study Conductor Obtaining Consent

Date

Printed Name of Study Conductor Obtaining Consent

Appendix II: Email Announcement

Email title:

Participants wanted for a usability study on the cultural impacts on multilingual website usability

Email Content:

Hello all:

My name is Mengze Zhou, and I am a second-year graduate student at SILS in UNC Chapel Hill. I am conducting my master project and would like to look for some participants to take part in my research study. This study will examine the impact of key culture markers (referring to “interface design elements and features that are prevalent, and possibly preferred, within a particular cultural group”) on the usability of the two multilingual website cases: MSN and Air China. The whole process consists of a pre-test questionnaire, a series of tasks that you will be asked to perform, a brief post-test interview, and a post-test questionnaire, which will take approximate 60-90 minutes to complete the whole process.

Data collected during the study will be associated with randomly assigned subject id number. All information contained on these and other forms are anonymous. Because the information is anonymous, we will not be able to withdraw it from the data pool. This project will be reviewed by the Office of Human Research Ethics of University of North Carolina at Chapel Hill.

If you are interested, please contact me at zmz1107@live.unc.edu and we can schedule the date for the study.

Thank you for your time.

Sincerely,

Mengze Zhou
014' MS in Information Science
School of Information and Library Science

Appendix III: Moderate Guide

Introduction

Moderator: Hello. Thank you for taking the time to participate in this usability test. The goal of today's session is to understand and examine the impact of key culture markers, referring to "interface design elements and features that are prevalent, and possibly preferred, within a particular cultural group", on the usability of the two multilingual website cases: MSN and Air China.

This session will take roughly sixty to ninety minutes, and you will need to complete four tasks, two on Chinese version of MSN/ Air China website and the rest two on English version of the other website. You may leave at any time without penalty during the procedure if you see fit. We will be recording audio and the activity on the screen during this test. You will not be personally identified from the data we collect, and the direct results from this test will remain confidential.

Please don't worry about trying to do things the "right" way during this test; we are testing the usability of websites, not you. Please be honest with your feedback and speak up to let me know what you are thinking and feeling. I may not always be able to respond to you immediately, but feel free to ask questions if you get stuck or confused.

[Hand participant the informed consent form]

With that out of the way, please review and sign the informed consent form before we proceed. If you have any questions or concerns, please let me know.

Any questions before we get started?

[Hand participant the pre-test questionnaire]

Please fill out this pre session questionnaire before we continue with the test.

(Website: Air China)

Task 1: Checking the flight schedule information

Moderator: In the first part of the test, I am going to ask you to complete two separate tasks using the English/Chinese version of website Air China.

[Hand participant the task list]

I'm going to give you a scenario for you to imagine. Here is a written copy in case you need to refer back to it:

You and your parents are planning to travel from Beijing, China to New York, USA for holiday. Your departure date is around the first week of April (any time between April 1 and April 7), and your stay at New York is about two weeks. Now you have chosen our Air China Airline and are about to check the flight schedule information for your round-trip from Beijing (PEK) to New York Kennedy Intl (JFK) on our website. After that, you can nail down the exact departure and return date that you feel the best, and then may ask your parent to make a booking.

As you perform the task, try to speak aloud what you are thinking. This will be a big help to us in understanding the usability of this website.

You may begin when you're ready.

Moderator and scribe should note:

Observe how the participant browses through the Air China homepage. Are there any hesitations or body language shifts worth noting?

Task is completed once the participant expresses that the task is complete. Allow the participant to browse, but do not let the participant continue on to make a real booking.

Post-task interview questions:

1. Did you have any difficulty with this task? If so, please explain.
2. Did you feel that you needed more information to complete this task? If yes, what additional information would you like?
3. Would you feel comfortable in general with the website so far?
4. Are there any costs to you at this stage of the process?

Task 2: Checking the airplane model information

Moderator: Now I'm going to give you another scenario for you to imagine for your next task. Here is a written copy in case you need to refer back to it:

[Hand participant the task list]

The departure flight from Beijing to New York that you feel satisfied with is Airbus 340, and your return flight from New York to Beijing that you've decided is Boeing 747. Now you want to know how the flight seats are distributed on those two airplane models, so you may tell your parent to make reservation for your preferred seats when they're booking. And also you want to know what kinds of entertainment gadgets are offered on those two flights.

As you perform the task, try to speak aloud what you are thinking. This will be a big help to us in understanding the usability of this website.

You may begin when you're ready.

Moderator and scribe should note:

Observe how the participant navigates through this task. Are there any hesitations or body language shifts worth noting?

Task is completed once the participant expresses that the task is complete. Allow the participant to browse, but do not let the participant continue on to make a real booking.

Post-task interview questions:

1. Did you have any difficulty with this task? If so, please explain.
2. Did you feel that you needed more information to complete this task? If yes, what additional information would you like?
3. Are there any costs to you at this stage of the process?

(Website: MSN)

Task 3: Checking the score of one NBA game

Moderator: In the next part of the test, I am going to ask you to complete two separate tasks using the English/Chinese version of website MSN.

[Hand participant the task list]

I'm going to give you a scenario for you to imagine. Here is a written copy in case you need to refer back to it:

You are a fan of Miami Heat and feel very excited about this new NBA season. Due to some reason, you missed its game against Chicago Bulls on March 9, and would like to check the score. Now you are at the web portal MSN for information (all the search should be conducted within the MSN website).

As you perform the task, try to speak aloud what you are thinking. This will be a big help to us in understanding the usability of this website.

You may begin when you're ready.

Moderator and scribe should note:

Observe how the participant navigates through this task. Are there any hesitations or body language shifts worth noting?

Task is completed once the participant thinks that s/he has found the information and expresses that the task is completed.

Post-task interview questions:

1. Did you have any difficulty with this task? If so, please explain.
2. Did you feel that you needed more information to complete this task? If yes, what additional information would you like?
3. Would you feel comfortable in general with the website so far?
4. Are there any costs to you at this stage of the process?

Task 4: Checking today's gold price

Moderator: Now I'm going to give you another scenario for you to imagine for your next task. Here is a written copy in case you need to refer back to it:

[Hand participant the task list]

You have bought some gold for investment, and you are concerned and want to know whether the gold price goes up or down today. You have come to this web portal MSN for information. (all the search should be conducted within the MSN website).

As you perform the task, try to speak aloud what you are thinking. This will be a big help to us in understanding the usability of this website.

You may begin when you're ready.

Moderator and scribe should note:

Observe how the participant navigates through this task. Are there any hesitations or body language shifts worth noting?

Task is completed once the participant thinks that s/he has found the information and expresses that the task is completed.

Post-task interview questions:

1. Did you have any difficulty with this task? If so, please explain.

2. Did you feel that you needed more information to complete this task? If yes, what additional information would you like?
3. Would you feel comfortable in general with the website so far?
4. Are there any costs to you at this stage of the process?

Moderator: You just completed the task section of the usability test. Before you enter the next two parts about your evaluation on cultural markers and cultural dimensions, I'd like you to fill out a brief questionnaire about your experience on the four tasks that you've done.

Appendix IV: Pre-test questionnaire

Thank you for participating in this usability test. Before we begin the test, please take a few minutes to answer the following questions about your background.

General Questions:

1. Are you male or female?
[Male; Female]
2. Have you participated in a usability test in the past six months?
[Yes; No]
3. Which of the following best describes your age?
[18 to 22; 23 to 29; 30 to 39; 40 to 59; 60 and older]
4. Are you currently a student?
[Yes, full time; Yes, part time; No]
5. What operating system do you usually use for personal computing?
[Mac; Windows; other]
6. What Internet browser(s) do you usually use?
[Internet Explorer; Firefox; Chrome; Safari; other]
7. Which language below can you understand?
[English; Chinese; both of them; neither of them]
8. Which culture below are you familiar with?
[English; Chinese; both of them; neither of them]

Domain Questions:

1. How often do you use online flight booking?
[Frequently; Sometimes; Rarely; Never]
2. If you have experience with online flight booking, which websites will you usually go to?
[Official websites of airlines; Online travel agencies; Other]
3. How often do you visit web portals for information?
[Frequently; Sometimes; Rarely; Never]
4. Which web portal(s) would you often like to go to?
[Yahoo!; About.com; MSN; Wikipedia; Other]
5. Have you ever been to websites with different multilingual versions
[Yes, often; Yes, sometimes; Yes, rarely; No]
6. Will you go to any websites for information with displayed language that you do not quite understand?
[Yes, often; Yes, sometimes; Yes, rarely; No]

Appendix V: Post-test questionnaire (for P1)

The following questionnaire is based on the USE Questionnaire but has been highly modified to fit this usability test. Each question is answered using a 7-point Likert-type rating scale (**1=Strongly Disagree, 7=Strongly Agree**). Category headings are not included in the questionnaire.

Ease of Use

- The Chinese version of _____ (filled by study conductor) is easy to use.

O	O	O	O	O	O	O
1	2	3	4	5	6	7

- The English version of _____ (filled by study conductor) is easy to use.

O	O	O	O	O	O	O
1	2	3	4	5	6	7

- The Chinese version of _____ (filled by study conductor) is user friendly.

O	O	O	O	O	O	O
1	2	3	4	5	6	7

- The English version of _____ (filled by study conductor) is user friendly.

O	O	O	O	O	O	O
1	2	3	4	5	6	7

Satisfaction

- I am satisfied with the Chinese version of _____ (filled by study conductor).

O	O	O	O	O	O	O
1	2	3	4	5	6	7

- I am satisfied with the English version of _____ (filled by study conductor).

O	O	O	O	O	O	O
1	2	3	4	5	6	7

Attractiveness

- The Chinese version of _____ (filled by study conductor) is visually attractive.

O	O	O	O	O	O	O
1	2	3	4	5	6	7

- The English version of _____ (filled by study conductor) is visually attractive.

O	O	O	O	O	O	O
1	2	3	4	5	6	7

- The Chinese version of _____ (filled by study conductor) has a clean and simple appearance.

0	0	0	0	0	0	0
1	2	3	4	5	6	7

- The English version of _____ (filled by study conductor) has a clean and simple appearance

0	0	0	0	0	0	0
1	2	3	4	5	6	7

Appendix VI: Post-test questionnaire (for P2)

In this part, evaluation and comparison will be focused on five major categories of cultural markers, which are: language/text, visual/graphic elements, colors, page layout and navigational function. A list of checkpoints will be used here, together with a scale of 1 to 5 is also used in rating the extent of user's perception on each of the five cultural markers: **1 = not perceptible** **2 = hardly perceptible**, **3 = perceptible to some extent**, **4 = clearly perceptible** and **5 = strongly perceptible**.

Checkpoints	Eng. ver. of _____	Chi. Ver. of _____	Notes/Comments
Language/text			
Text density			
Text size (big?)			
Text style integrity			
Language integrity			
Content integrity			
Visual/graphic elements			
Logo/brand			
Image density			
Image size (big?)			
Image resolution			
Video display			
Animation			
Colors			
Color intensity			
Color relations (close?)			
Color diversity			

Page layout			
Page length			
Centered or not			
Regular or not			
Navigational function			
Menu-bar intensity			
Button intensity			
Link intensity			
Flexibility (easy to go to different levels?)			

Appendix VII: Semi-structured interview (for P3)

Relying on the literature frameworks that have been proposed by Hofstede, the cultural dimensions that are used for verification are: Power distance (PD), Collectivism vs. Individualism (IND), Femininity vs. Masculinity (MAS), Uncertainty Avoidance (UAI), and Long and Short Term Orientation (LTO).

The five cultural dimensions along with its explanation and effects on web design are explained by the study conductor, and will also be printed out and handed out.

Power Distance (PD)

Power distance refers to the extent to which less powerful members expect and accept unequal power distribution within a culture.

- Hierarchies in website structure: tall vs. shallow.
- Focus on expertise, authority, certifications, official stamps, or logos: strong vs. weak.
- Importance of restrictions or barriers to access: frequent restrictions on users vs. transparent.

Individualism vs. Collectivism (IC)

Individualism in cultures implies loose ties; everyone is expected to look after one's self or immediate family but no one else. Collectivism implies that people are integrated from birth into strong, cohesive groups that protect them in exchange for unquestioning loyalty.

- Argumentative speech tolerance vs. official slogans and subdued controversy
- Emphasis on change: new and unique vs. tradition

Masculinity vs. Femininity (MAS)

Hofstede focuses on the traditional assignment to masculine roles of assertiveness, competition, and toughness, and to feminine roles of orientation to home and children, people, and tenderness.

- High-masculinity cultures would focus on:
 - ✓ Traditional gender/family/age distinctions
 - ✓ Work tasks, roles, and mastery
 - ✓ Navigation oriented to exploration and control
 - ✓ Graphics, sound, and animation used for utilitarian purposes
- Feminine cultures would emphasize the following:
 - ✓ Attention gained through poetry, visual aesthetics, and appeals to unifying values

Uncertainty Avoidance (UA)

Cultures vary in their avoidance of uncertainty, creating different rituals and having different values regarding formality, punctuality, legal-religious-social requirements, and tolerance for ambiguity.

- High UA would emphasize on:
 - ✓ Simplicity, with clear metaphors, limited choices, and restricted amounts of data.
 - ✓ Navigation schemes intended to prevent users from becoming lost.
 - ✓ Redundant cues (color, typography, sound, etc.) to reduce ambiguity.
- Low UA would focus on:
 - ✓ Complexity with maximal content and choices
 - ✓ Less control of navigation

Long- vs. Short-Term Time Orientation (LTO)

Hofstede concluded that Asian countries are oriented to practice and the search for virtuous behavior while Western countries are oriented to belief and the search for truth.

- High LTO would emphasize the following aspects:
 - ✓ Content focused on practice and practical value
 - ✓ Patience in achieving results and goals
- Low LTO countries would emphasize the contrary:
 - ✓ Content focused on truth and certainty of beliefs
 - ✓ Desire for immediate results and achievement of goals

The explanation above is provided to each participant to give them a better understanding on Hofstede's five cultural dimensions, which is cited from: A. Marcus and E.W. Gould. Cultural dimensions and global web user-interface design: What? so what? now what. Proceedings of the 6th Conference on Human Factors and the Web, 1–15, 2000.