

UNIVERSITY NAME

DOCTORAL THESIS

Thesis Title

Author:
John SMITH

Supervisor:
Dr. James SMITH

*A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy*

in the

Research Group Name
Department or School Name

May 12, 2018

Declaration of Authorship

I, John SMITH, declare that this thesis titled, "Thesis Title" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."

Dave Barry

UNIVERSITY NAME

Abstract

Faculty Name
Department or School Name

Doctor of Philosophy

Thesis Title

by John SMITH

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Acknowledgements

The acknowledgments and the people to thank go here, don't forget to include your project advisor...

Contents

Declaration of Authorship	iii
Abstract	vii
Acknowledgements	ix
1 Introduction	1
1.1 Background	1
1.2 Global Website design	1
1.3 Tetra Pak	2
1.4 Purpose	2
1.5 Project scope	2
2 Theory	3
2.1 Cultural differences in Perception	3
2.2 User Experience	3
2.3 F-shaped Pattern	4
2.4 Perception in asia (f-shaped pattern.)	4
2.5 Usability Metrics	4
2.5.1 Performance Metrics	4
2.5.2 Self-Reported Metrics	6
2.6 Usability Testing	6
2.7 Colour and Culture	6
2.8 Trends	6
2.9 Culture and Usability	7
2.10 Great Firewall of China	7
2.11 Asynchronous	7
2.12 AWS - Amazon Web Services	7
2.12.1 EC2	7
2.12.2 Auto scaling	7
2.12.3 Load balancing	7
2.12.4 RDS	8
2.12.5 S3	8
2.12.6 Elastic Beanstalk	8
2.13 React-Redux	8
2.13.1 Redux-saga	8
2.14 MySql database	8
2.15 API	9
2.16 Statistical significance	9
2.17 waterfall method	9
3 Methodology	11
3.1 Structure	11

4 Phase 1 - Investigation	13
4.1 Work process	13
4.2 Chinese vs Western websites comparison	13
4.2.1 BBC	13
4.2.2 QQ	14
4.2.3 Taobao and Ebay	14
4.2.4 Analyses of Ctrip	15
4.3 Analysis	16
4.3.1 Work based on analysis	18
5 Phase 2 - Prototyping	19
5.1 Work process	19
5.2 Prototypes	19
5.2.1 Low-fi Prototype	19
5.2.2 High-fi Prototype	19
5.2.3 Questionnaire	20
5.3 Analysis of results	20
6 Pilot study	21
6.1 Method	21
6.2 Results	21
6.3 Discussion	23
6.4 Conclusion	24
7 Phase 3 - Building the Interfaces	25
7.1 Work process	25
7.1.1 Front-End	25
Homepage	26
BBC and QQ	28
Questionnaire	28
7.1.2 Database	29
7.1.3 Api	30
7.1.4 Hosting AWS	30
7.1.5 Beta-tests	31
7.1.6 Launch	31
7.1.7 How the test was conducted	32
8 Phase 4 - Analyzing Data	33
8.1 Work process	33
8.2 Analysis	34
8.3 Questionnaire	34
9 Results	35
9.1 BBC English	35
9.2 BBC Chinese	35
9.3 Differences between Chinese and English BBC users	36
9.4 QQ English	37
9.5 QQ Chinese	38
9.6 Differences between Chinese and English QQ users	38
9.7 Questionnaire Results	40
10 Discussion	41

11 Conclusion	43
11.1 Limitations	44
11.2 Future work	44
A Appendix A	45
Bibliography	49

List of Figures

3.1 Homepage	11
4.1 QQ.com	14
4.2 QQ's Menu bar	14
4.3 ebay	15
4.4 Ebay's menu bar	15
4.5 Taobao	16
4.6 Taobao' menu bar	16
4.7 Chinese version of Ctrip	17
4.8 English version of Ctrip	18
6.1 BBC pilot study results	22
6.2 QQ pilot study results	23
7.1 Homepage	27
7.2 Homepage	27
7.3 Users View	28
7.4 SUS	29
7.5 Database Schema	30
9.1 BBC Results	37
9.2 QQ Results	39
9.3 Questionnaire Results	40
A.1 QQ	45
A.2 BBC	46
A.3 Done	47

List of Tables

Chapter 1

Introduction

This master thesis has been completed for the Department of Design Sciences at Lund University in collaboration with the company Tetra pak. The first chapter will introduce the background, purpose and scope.

1.1 Background

Working in Shanghai for a summer, I noticed that there was a big difference in how Chinese websites were designed compared to western sites. I found the Chinese sites very overwhelming in information density. I assumed that this was because of a simple difference in cultural trends. Later I came across some cross-cultural research articles proving that there is a difference in how people from western and eastern societies perceive information. As such, I wanted to investigate if this also influences how the Chinese web pages are designed? Together with Tetra pak, this thesis project is created to examine the differences in web design in eastern and western cultures and to find out if a global interface for both cultures can be created.

1.2 Global Website design

Designing websites for a global market is something that is becoming increasingly necessary for large multinational companies to do. In spite of the need and popularity of global designed websites there is a lack of research in this field for user experience. Many companies simply try to launch their local product globally and hope the design works everywhere. Some companies, conversely, simply create a new product for the different market without researching if this is necessary. It is evident that websites look different in cultures all over the world. But how much of this difference is because of cultural trends? Are parts of these websites designed after a difference in how individuals from different cultures perceive information? How well can users from different cultures perform on different web elements?

There is a surprising lack of research done on this subject. Despite this, however, companies over the world continue to spend millions attempting to build information products for cultures without relevant knowledge about how users from that culture perceive information. There are quite a few assumptions that are commonly made about users in different cultures. The first and most common assumption is that all users process information the same way. Although this assumption has been found false, most people believe that the differences in web layout are mostly due to a difference in language or simply trends. The second assumption is that because eastern websites are designed differently, easterners prefer this type of design. There are quite a lot of ambiguities when examining web pages and web

apps in different cultures and hopefully, this thesis will resolve some of these uncertainties.

1.3 Tetra Pak

Tetra Pak is one of the world's leading food processing and packaging solutions company. With more than 24,000 employees around the world they are one of the biggest actors when it comes to food safety. Tetra pak provide everything from factory processing, packaging, automated plant solutions to web based systems that helps their users control their factories.

1.4 Purpose

The purpose of this thesis is to research the differences between western and eastern website usage, and how differences in perception plays a role in this. To do this we will try to answer the following questions:

1. Are differences in interface design due to differences in information processing styles or trends?
2. Do different processing styles in Western (analytical) versus Chinese (holistic) users significantly affect performance on different interfaces?
3. Can one Global interface be created, or should web designers focus on creating separate user interfaces for different cultures?

This thesis is done in partnership with Tetra pak since they currently have development ongoing in many eastern countries, such as China. Tetra pak is interested in trying to learn more about user-experience in these eastern countries to help them improve their products in these countries. Using Tetra pak resources from China allowed this thesis to be conducted and is a reason for limiting the testing to Chinese and Swedish users.

1.5 Project scope

The duration of this project is limited to 20 weeks that, once completed, will deliver answers to the questions defined in the purpose above. This time limit does create some restrictions for the study.

1. Majority of time had to be spent on the development phase since this is the most crucial part to be able to gather any data at all.
2. Once a certain time had past no more data could be gathered from users. This for the thesis to be finished in time. Even if this could possibly lead to more statistically sound results.
3. Focus had to be put on a select few user behaviours. Once again because of the time limitations.

Chapter 2

Theory

2.1 Cultural differences in Perception

Individuals hailing from different cultures have been found to process the same, objective information differently. This variation in perception is theorized to be due to how different cultures encourage and train individuals to process information differently. In particular, the influence of culture on information processing appears most potent when considering how easterners (e.g., Chinese, Japanese, etc.) and those of western cultures (e.g., American, Swedish, etc.) perceive information differently [8]. Specifically, research suggests westerners, who tend to be individualists, interpret content analytically while easterners, who tend to hold collectivists cultures, usually process information more holistically)[9]. The difference between western analytical thinking and eastern holistic thinking is that holistic thinking involves perceiving information through the interconnectedness of the perceptual stimuli, while analytical thinking involves the interpretation of objects independently, in isolation of surrounding cues.[12]

Essentially, easterners appear to take in more content when viewing an object than westerners do. On eye tracking studies, east Asians eyes tended to scan through an entire image, bouncing around to intake all of the information, even background content. Conversely, eye tracking heat maps indicate that western subjects did the opposite, instead of taking account of all the objects in an image, westerners tended to look at the biggest most central object in great detail. That is, westerners could recall more specific details about one central object and their eyes lingered on one item longer, while easterners look at everything in an image and their eyes bounce around. As such, holistic processing "involves an orientation to the context or fields as a whole" [10] whereas analytical processing "involves a detachment of the object from its context, a tendency to focus on attributes of the object" [12]). These findings suggest that eastern collectivists, such as individuals from China, will be more likely in-take more information content than westerners and to use non-central cues for information processing. Conversely, it suggests that western analytical thinkers may prefer to focus on fewer objects and may tend to use only central information.

2.2 User Experience

User Experience (UX) is an expression popularised by Donald Norman and Jakob Nielsen in "The design of everyday things" (länska på rätt sätt)(Källa!). Norman and Nielsen define User Experience in the following way "True user experience goes far beyond giving customers what they say they want, or providing checklist features. In order to achieve high-quality user experience in a company's offerings there must

be a seamless merging of the services of multiple disciplines, including engineering, marketing, graphical and industrial design, and interface design" (Norman and Nielsen, 2016). (latex quota på rätt sätt).

The term "User Experience" is a widely applied term in web design and it is associated with a multitude of definitions. The International Organization for Standardization (ISO) has also attempted to define the term. Specifically, the organization, in ISO 9241-210, defines UX as "A person's perceptions and responses that result from the use and/or anticipated use of a product, system or service." Further, the ISO states that "UX includes all users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use". The standard also states that "UX is a consequence of functionality, system performance, interactive behaviour and assistive capabilities of the interactive system, the user's internal and physical state resulting from prior brand image, presentation, experiences, attitudes, skills and personality, and the context of use"

2.3 F-shaped Pattern

The F-shaped pattern refers to findings made in a study on user eye movements [15] (find correct article for f-shaped pattern and cite it here as well). This pattern has been dubbed the "F-shaped pattern" since the study found that users often scan through pages starting with a horizontal movement, usually across the upper part of the content area. Users then tend to read across in a second horizontal movement further down on the page that typically spans a shorter area. Lastly, users scan the content's left side in a vertical movement. When measuring users' eye gazing as a heat map, this creates a pattern that resembles an F. Web developers, either knowingly or unknowingly, often design their websites according to this pattern. The F-shaped pattern is not an absolute law and several other scanning patterns exist, but the F-shaped pattern remains the most prevalent in western cultures. [14] If a developer designs a page without knowledge about this pattern, they risk putting important information in places where users might miss it. (LÄGG TILL BILD)

2.4 Perception in asia (f-shaped pattern.)

2.5 Usability Metrics

There are several different types of metrics that can be used to measure the usability of a prototype or product. Among these metrics are performance metrics, Issues-based metrics, self-reported metrics, behavioral metrics, comparative metrics, and etc [23]. For this project, we have decided to focus on performance metrics and self-reported metrics. Usability metrics are powerful tools that are generally under utilized [11].

2.5.1 Performance Metrics

Performance metrics can be used to measure a user's behavior when interacting with a product. In this project, the performance metric data will automatically be gathered and sent to a database. This data will then be further analyzed to gain a deeper understanding of how the users engage with the product.

There are 5 basic performance metrics which include: [23]

- Task success
- Time-on-task
- Errors
- Efficiency
- Learnability

In order to measure the task rate success metric, the required task must be clearly defined and have an unambiguous goal. For example, "send an email to person x" reflects a well structured task, thus, task success rate can be accurately measured given the well-defined objective. Conversely, the task, "research budget car brands" does not offer clear guidelines and can be responded to in multitude of ways due to its ill-defined objective and, as such, would not be suitable for measuring task success. There are two types of task success forms. The first is represented as a binary measure; either the user completes or fails to complete the task [23]. The second type is continuous rather than categorical as it measures the level of success. This metric is particularly useful if the given task can be partially completed. One example of a task that would benefit from partial success rate measurements would be if we ask users to open a specific video on YouTube, but the user opens the wrong video. The user would still be somewhat correct as she correctly navigated to YouTube, but failed to select the correct video, yielding a partial success. The simplest way of measuring success levels is to assign numeric values to the study subject's performance (i.e., this might range from 1 to 10, where 5 represents that the user completed half of the task).

There are several ways in which users' can fail to accomplish their tasks. Users may, for instance, incorrectly assume the task has been completed when, in fact, it is only partially complete. Further, users may give up on trying to solve the task out of frustration or users may falsely believe they successfully finished the task when, in actuality, they performed the wrong task. This data can be invaluable to the process of uncovering how well users understand a given system.

Time-on-task is a simple measure; it simply logs the time users' spent to complete or fail the task at hand.

Errors in this case do not refer to programmatic errors, but to mistakes made by the user. One example of an error is when users select an incorrect tab before finding the correct one. In this example, every wrong path and/or click used in performing the task, except the optimal one, is logged as an error. Error measurements can assist developers in gauging how well a user understands the website as well as how intuitive the website is for a first time user.

Efficiency is analogous to the Time-on-task measure, but it can also be measured by how many steps users took to complete the task. It is important to note that efficiency should only be measured on successful tasks [23].

Learnability represents the degree to which a user becomes more efficient at using a product over time. Essentially, the amount of time reduced in completing a task

from the first iteration to the second or third provides an indication of how well the user learned to interact with the product.

2.5.2 Self-Reported Metrics

Self-reported metrics are employed when directly asking users to describe their experience engaging with the product. One means of obtaining self-reports is through surveys. A common method for doing this is by employing the System Usability Scale (SUS) created by John Brooke [23] [5]. SUS is a survey containing 10 questions with a scale from 1-5, where 5 corresponds to "strongly agree" and 1 reflects "strongly disagree", for each question (See appendix X). SUS is a metric tool, used over 22 years, that has been proven to be a streamlined and robust tool for gauging usability [5].

2.6 Usability Testing

Use this to describe the gathering of data needed to find most faults in a interface [23].

2.7 Colour and Culture

Different cultures have always preferred distinct colour schemes, and this variation has also effected the degree to which a user trusts or likes a website. Not all people prefer the same colour layouts and a study conducted by Cyr, Head, and Larrios [6] indicates that this colour bias may be rooted in culture. The study suggests that the colour schema a website uses influences the trust and likeability of the web page. The study further indicates that individuals hailing from different cultural backgrounds tend to select colours associated with their culture. These cultural factors must be accounted for when designing websites for any given culture as colour schema impacts how well users like and engage with the website. Selecting colours that users from a certain culture feel more comfortable with can be vital to enhancing the user's experience with the site.

2.8 Trends

Trends, simply something that is popular in a particular moment, are an ever-present phenomenon. However, simply because a certain behaviour or style is trending does not necessarily indicate that it is the most optimal way to perform an action; on the contrary, usually the opposite is true. Comparing design trends to actual usability in this research means that we will examine if there is any actual underlying data that supports the trend from a usability perspective. This can have two outcomes: either the trend has grown because it more closely caters to how users use the respective products or the trend is a by-product from how designs were previously created. One example of this is that we load more information than necessary on to a page because we have previously done so. The reason we started doing this was due to slow internet speeds which lead to large loading times when clicking through a page. As such, now, even if the internet speed is quick and we do not have to load all information to a page we still do it since both users and developers have become accustomed to this pattern. [Source for this?]

2.9 Culture and Usability

2.10 Great Firewall of China

The Great Firewall of China (GFC) is a combination of laws and technologies the Chinese government uses to domestically regulate the internet. Examples of services blocked by the GFC are Google, Facebook, and Youtube. The GFC also artificially causes traffic from abroad to be significantly slower than applications hosted in China. Hosting an application on a server in China requires a specific IPC license from the Chinese government and getting one is a long and slow bureaucratic process. The sort of algorithms that are used by GFC are largely unknown and can be challenging to circumvent [20].

2.11 Asynchronous

Asynchronous programming refers to the task of making several data processors run in parallel to each other usually without impacting one another. Asynchronous parallel processes are often called threads. One example of this would be one thread working on reacting to a user's request and supplying that user with the correct information. Simultaneously, another thread, not visible to the user, is saving all the user's actions and sending them to a server.

2.12 AWS - Amazon Web Services

Amazon Web Services (AWS) is the world's largest provider of web-hosting. Amazon allows their customers to easily host applications globally and the company provides several features to assist their customers with this task [1].

2.12.1 EC2

EC2 (Elastic Cloud Compute) is a basic web server service that AWS offers. EC2 allows customers to set up a virtual server with different amounts of CPU, Memory and etc. These servers can be set up on several AWS locations across the world. This server can be customised to run an operating system of the customer's choice, the most common of which are Linux and Windows. [3]

2.12.2 Auto scaling

Auto scaling is a feature provided by AWS that automatically scales up the server in case of increased traffic. This means if a application has a large amount of traffic on a server the auto scaling functionality creates an extra server that can handle user requests. Auto scaling also allows for automatic scale down in case of low traffic. [3]

2.12.3 Load balancing

Load balancing is a feature from AWS that automatically balances the load of the EC2 instances. If a user has three EC2 instances, the load balancing will make sure that the workload is shared by all EC2 instances. This helps to prevent one instance from overloading. [3]

2.12.4 RDS

Relational database service (RDS) is a database service provided by AWS. RDS lets customers set up a database of their choice and host it on AWS servers. Customers can set this database up on several locations across the world and configure it to suit their application. [3]

2.12.5 S3

S3 is an AWS feature that allows for object storage in the cloud. S3 allows the user to store anything he deems fit - this be everything from files and Images to code repositories. Images that are used on websites can stored on S3 and then downloaded to the website when the user opens it, this is a common way to handle images in web sites and applications. [3]

2.12.6 Elastic Beanstalk

Elastic Beanstalk also called EB is a feature provided by AWS that automatically sets up a instance complete environment with auto scaling, load balancing, Relational database and EC2 instances. [3]

2.13 React-Redux

React is a front-end JavaScript library developed by Facebook. React is based on the user building and reusing components. This allows for very structured and highly scalable code.

Handling data-flow in a react application can be very tricky, this is where redux comes in. Redux is a JavaScript library that allows for structuring and handling of a web application's data flow in a structured way. React and Redux are so commonly used together that libraries combining them have been made. React-Redux is the most popular use of these libraries and they work very well together to allow scalable and reusable code. [17] [18]

2.13.1 Redux-saga

Redux-saga is a javascript library that is made to handle a applications asynchronous tasks. Redux-sagas is often used for data feching and posting. It can also be used for other asyncronius tasks. Sagas handle asynchronous tasks without the user getting impacted at all by what goes on in the background. [19]

2.14 MySql database

MySql is a version of the database query language SQL. SQL has been used since 1981 and is used to set-up, save and get information from a database. MySql is free to use and has a public license. Mysql is a language that is both simple to use and quite powerful. Setting up inputting and getting data from an SQL database can be done through only a few lines of code.

2.15 API

Api (application programming interface) is a interface between the front-end and server. A api allows the application to communicate with functions and servers outside the internal environment. Examples of these are databases, other servers and other api's. An api allows for clearer communication between different actors on the web.

2.16 Statistical significance

Statistical significance is a term that commonly appears in scientific research and it refers to the likelihood that two or more variables are related beyond just random chance. If the relationship between two entities usually occurs by random chance, for example, the relationship between the size of your house and the amount of food you eat, this connection between the variables would not be statistically significant. Conversely, the relationship between the size of an individual's house and the amount of property tax that individual pays would, most likely, be statistically significant as the variables are often related beyond random chance. In order to test for significance, scientists and researchers employ hypothesis testing to analyse if the data is significant. Further, researchers employ tests (such as t-tests) to find the p-value of their data, which reflects the probability that the results could be explained by randomness alone. Generally, a rigid and scientifically solid study has the p-value threshold set at 5% (the threshold I use in my studies), which indicates that the probability of this event or difference occurring based on pure random chance is less than 1 in 20. In my research, I explore if the differences between means (Chinese users vs. western users) are large enough to be non-random, suggesting that there is an actual definitive distinction between the two groups. [2]

2.17 waterfall method

Chapter 3

Methodology

This chapter aims to present the methodology, layout used in this report and the work process of this Thesis.

3.1 Structure

This project followed a waterfall based approach where each smaller element in the process was done iteratively. The waterfall based approach suited this project since each phase had to be completely finished before moving on to the next part. Each one of these phases, on the other hand, followed a more iterative agile process. One example is the development part of the project, this phase could not start until the high-fi was completely done and tested. Two of the phases were done iteratively, on the other hand (see 3.1).



FIGURE 3.1: Flow Chart showing the different phases. Arrows indicate iterative phases.

The phases that can be seen in the figure 3.1, represents the different steps of the work process. Each new phase was dependent on the results of the last phase and at the end of each phase what was needed to be done and used in the next phase was decided. Since each phase can be seen as a smaller project they are all divided into a separate chapter in this thesis. Each chapter explains the work process during the phase. If relevant, results for that phase is explained. Lastly, once again if necessary for the phase, results where analysed and changes will be described in the chapter.

Following the phase chapters, results of the actual test are shown in the "Results" chapter, this includes numbers, graphs and calculations done etc. These results are then discussed in the following chapter "Discussion". In the "Conclusion" chapter i will explain my conclusion, what limitations exist and future research necessary.

Chapter 4

Phase 1 - Investigation

The aim of the investigation phase is to obtain a more nuanced understanding regarding the differences in perception and navigation of western versus Chinese websites. This information will facilitate the decision making process concerning which features to keep, which features need to be analysed, and what usability metrics should be implemented in the study.

4.1 Work process

As an initial step, Chinese websites along with their western equivalents were identified and curated. After obtaining a corpus of websites from China and western countries (e.g., United Kingdom), the key design differences between these respective sites were documented. Once the primary design variances were explored, this information was used to decide which design patterns should be tested in the study. Finally, the specific metrics to measure the results of these different designs were concluded.

4.2 Chinese vs Western websites comparison

In reviewing the websites, it was evident that Chinese pages differed significantly from western counterparts. The differences between sites extend from the look and feel to the UX design. A few of China's most popular browsing sites will be analysed and compared to similar western counterparts to further investigate these design distinctions.

4.2.1 BBC

BBC is a very popular western news site. The design is quite westerly influenced in terms of information density and design. As can be seen in the fig... the site follows the f-shaped pattern quite heavily. Each news section has a large news figure in the top left corner and two rows. Just looking at the large picture and the first row then glazing down to the next news section would have the users quite naturally follow the f-shaped pattern (fig med röd färg). The site itself is quite information sparse with large images taking most of the space. Without hovering over any content on a standard computer screen, there are roughly 48 clickable elements.

4.2.2 QQ

QQ is one of the most visited websites in China (see fig 4.1). [22] [21] QQ, like many other Chinese websites, does not focus on one thing but has multiple different functions. Some of the functions that QQ supports are: instant messaging, online games, music, shopping, microblogging, news, movies, group chat software, and etc. On the QQ homepage, users are greeted by the site's news page, which is highly information dense. Without hoovering over any content on a standard computer screen, there are roughly 147 clickable elements. In contrast, BBC's homepage [4], which is considered fairly information dense by western standards, has only 48 clickable elements on its homepage. This means that with only counting clickable elements, QQ is over 3 times more information dense than BBC.



FIGURE 4.1: QQ's homepage which provide which is mostly used for news.

One element that is notably common on Chinese websites, including QQ, is the menu bar design (see fig 4.2). On the QQ page, the menu bar contains two rows with a total of 40 clickable options - this format of menus is typical in China and is shared by multiple other Chinese sites.



FIGURE 4.2: A close-up of the menu bar used at QQ.

4.2.3 Taobao and Ebay

Taobao, which provides services akin to America's Ebay, is one of the world's biggest e-commerce platforms. Both Taobao and Ebay are shopping websites where users can purchase nearly any product they need, both from retailers and from other consumers. However, although these sites are similar in service, the design and user-experience focus on these sites differ significantly. Ebay, for instance, boasts a sleek design, employing dark-themed colors and contains only 20 clickable elements on the home page (see fig:4.3). Conversely, the Taobao page contains more colors, employs a brighter theme, and incorporates more clickable elements compared to its western counterpart. Another notable distinction between the websites is that

Ebay has an expanding menu bar containing roughly 6-10 clickable elements while Taobao's menu contains much more (see fig:4.4).

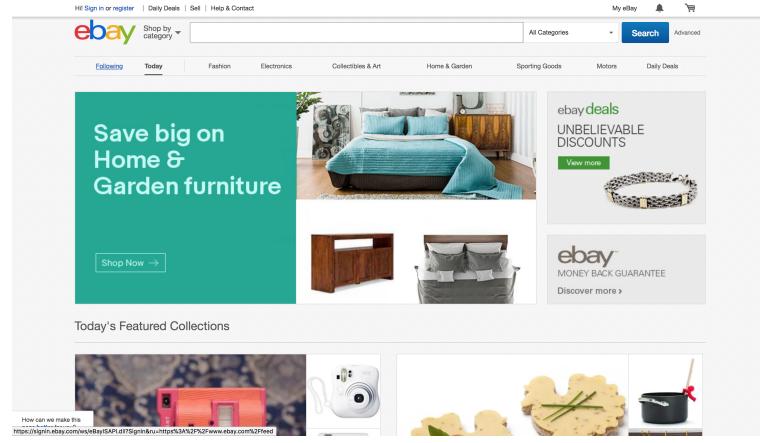


FIGURE 4.3: Ebay a popular American online shopping site

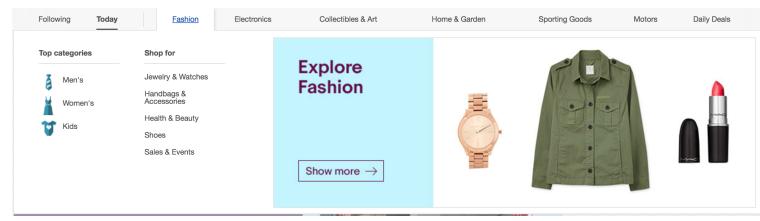


FIGURE 4.4: Expanding the menu on Ebay.

In examining the Chinese version of Taobao, the difference in information density is evident (see fig:4.5). The main page, for example, hosts roughly 49 clickable elements. Additionally, the menu items on the screen's right side is expandable, displaying between 55 to 88 clickable links and elements see fig:4.6). This amounts to about eight times the amount of clickable elements on Ebay. Further, while Taobao employs strong color themes (e.g., red, purple, orange, and blue), Ebay generally employs muted colors (e.g., grey) allowing for the products to be the center of focus.

4.2.4 Analyses of Ctrip

Interestingly, the layout on many Chinese websites change significantly when the language is changed. For instance, the layout of Ctrip, a common travel site used for booking hotels and flights in China, becomes very different when users select English for the site (see fig: 4.7 versus for the Chinese version (see fig: 4.8 for English version)). With the exception of the brand and name of the website, it is difficult to tell that it is actually the same website .

The main difference documented between the sites is content density. The Chinese version of Ctrip has a lot more content in a small area compared to the English-translated version of the site. Counting clickable elements without hovering over anything, 40 clickable elements were found on the Chinese version compared to only 26 clickable elements on the English version. When using the Chinese site, all links open a separate window instead of a second menu or tab - a common phenomena found on Chinese sites.



FIGURE 4.5: Taobao a popular Chinese online shopping site

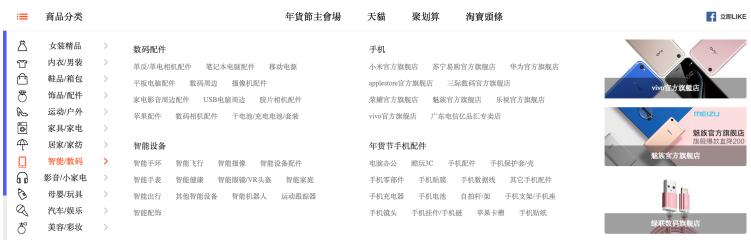


FIGURE 4.6: Expanding the menu on Taobao.

4.3 Analysis

Looking through the websites we can identify several design features (outside of the language differences) that differ in Chinese and western websites.

These are:

- High information density
- Colors
- Ad content
- Navigation

The main factor of variance across the sites is information density. Chinese websites have significantly higher information density compared to their western counterparts. As such, information density is one feature that will be closely examined in this research. Colors and navigation features will be explored, but not prioritized. These aforementioned features will be included to aid in understanding the look and feel of the sites, rather than being the primary objectives of investigation. Additionally, since Chinese sites contain a higher ratio of ad content compared to western sites. There are several theories accounting for why Chinese sites are more information dense. Some of these hypotheses are: cultural/trends, historical, holistic vs analytic perceptions and language. Due to limitations concerning Chinese language and knowledge of the history and culture, we will mainly examine trends and perception.

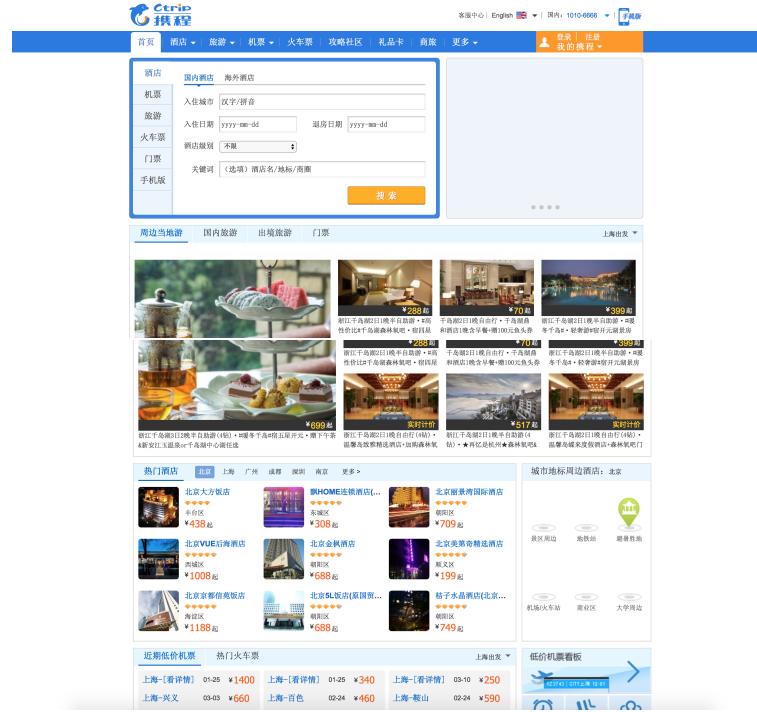


FIGURE 4.7: The Chinese version of the travel website Ctrip.

For the purposes of this research, different interfaces were created for the purpose of investigating trends and perception differences between sites. One page is western inspired and another was created with inspiration from Chinese designs. To ensure that the interfaces maintain authentic Chinese and western designs, professional UX-designers from Sweden and China assisted in the development of some prototypes. These prototypes were then tested on both users with Swedish and Chinese heritages respectively. Finally, working interfaces, capable of measuring what actions users take in responding to certain tasks, was developed given the feedback from these prototypes. Main measurements that will be used are task-success, time-on-task and a modified System usability scale. [5]

Four interfaces will be created, with two following a western design and the other two using a Chinese layout. The first interface takes inspiration from the QQ and BBC news site home pages. The goal of implementing these interfaces is to explore how fluidly users from different cultural backgrounds can navigate sites containing high information density (e.g., copious amounts of images and texts). Both interfaces contain roughly equivalent levels of material and clickable elements; the primary difference is that the western site will be longer, forcing users to scroll down the page. Additionally, some of the information will be mapped in sub-menus using natural mapping for the western site [13]. Conversely, the Chinese inspired site will provide most of the material directly on the screen for users to view without any nested menus.

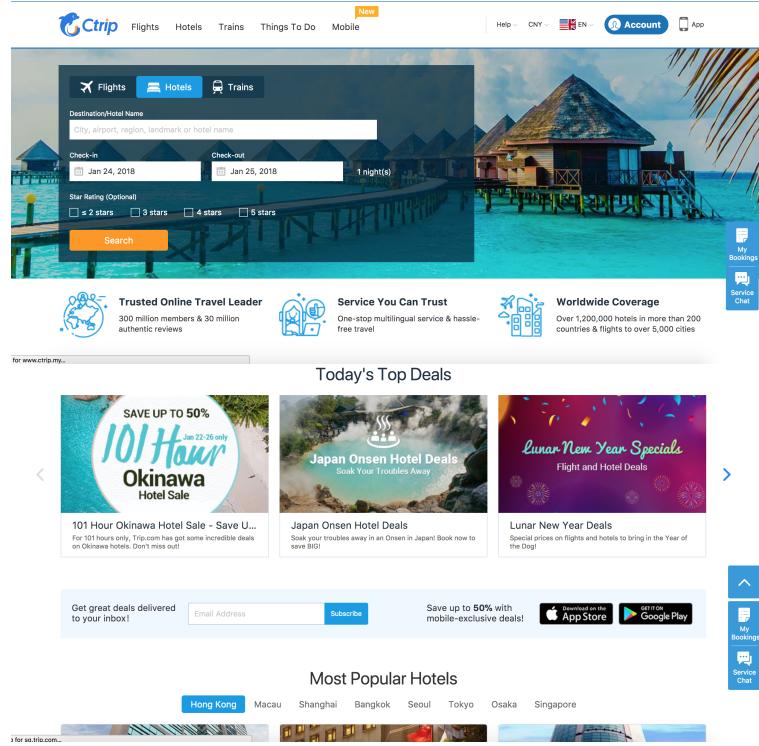


FIGURE 4.8: The English version of the travel website Ctrip.

4.3.1 Work based on analysis

Choosing the ux questions: In [12] has shown that perception differs in western and eastern cultures. [7] further indicates that this perception difference holds true in the case of people observing websites. Specifically, users using analytical processing follow the F-shaped pattern when browsing sites [15]. Holistic thinkers, conversely, do not follow the F-shaped pattern when browsing through a website. [7] In light of these past studies, it will be interesting to explore how these perception differences will influence users' abilities to navigate and perceive web pages. To investigate this question, we select elements both in accordance to the F-shaped pattern and elements outside of this pattern. By testing the performance on both analytical and holistic thinkers, we should get an indication of the differences and how well people follow the F-shaped pattern when looking for specific elements. The test will be unsupervised, meaning that we will have to get a larger test audience in order to obtain significant results. To test variances in perception, we will create tests for the sites BBC and QQ. On these pages, we will ask test subjects to find elements following an F-shaped pattern and elements not following this pattern.

Chapter 5

Phase 2 - Prototyping

The goal of Phase 2 is to quickly create prototypes for our design that can be used in a pilot test to ensure that the real test will work as intended before it is created.

5.1 Work process

The Low-fi prototype was a simple sketch made with paper and pen. Since the prototype are modelled after an existing website this prototype is focused on the parts needed for testing the page and not the design of the web pages themselves. The low-fi prototype was then discussed with a UX-expert and improved upon.

Two High-fi prototypes was made from the online news site BBC [4] and QQ [16]. These prototypes were directly modelled from the websites and then the corresponding logos were removed. Both the high-fi prototypes are then translated to English respectively Chinese using Google translate. This translated text was then looked over by the Tetra pak supervisor who is a native Chinese speaker to ensure that the translation was correct. Test questions and tasks was created for both the BBC and the QQ site in English. These questions were also translated to Chinese for the Chinese users.

A questionnaire modelled after the Sus (system usability scale) was then created. The purpose of this questionnaire is to gage what the user feel about using the websites. Many of the original questions did not suit this purpose and was therefore either removed or changed.

5.2 Prototypes

5.2.1 Low-fi Prototype

The Low-fi Prototype can be seen in the following figure (fig:.....). As can bee seen in the images only the testing part and not the design of the website was created. The improvements from this prototype can be seen in the high-fi prototype.

5.2.2 High-fi Prototype

The High-fi prototypes can be seen in the following figures: QQ (CITE to QQ image), BBC (SITE TO BBC IMAGE). The test parts can be seen in the following image(ref to that figure)

5.2.3 Questionnaire

The Sus inspired questionnaire contained the following questions:

1. I liked the design of the site.
2. The design of this site was similar to other news sites.
3. I think that I would like to use this site frequently.
4. I found the site unnecessarily complex.
5. I thought the site was easy to use.
6. I found the various functions in this site were well integrated.
7. I thought there was too much inconsistency in this site.
8. I would imagine that most people would learn to use this site very quickly.
9. I found the site very cumbersome to use.
10. I felt very confident using the site.
11. I thought that the amount of information on this site was.

For each of these questions the user could give it a rating from 1 "Strongly disagree" to 5 "Strongly agree". The last question the user instead got the choices 1 "Too sparse" to 5 "Too much".

5.3 Analysis of results

The main purpose of creating and testing these high-fi prototypes was to see how well the site would work when translating to another language. The site got a very different look after being translated, partly because the Chinese language produces smaller sentences. The Chinese characters also give a very different impression since the design of the characters and general lack of fonts when using. The Chinese QQ site became quite a bit longer with the translated English text. And many sentences had to be spread out over two rows. For the translations made from Chinese to English, many news had to be slightly changed to make logical sense. The translator handled English to Chinese quite a lot better and only small changes needed to be done. These changes in between version should not have any effect on the testing results since what is being tested are how the users handle information density and placement not the actual content of the news.

The questions/tasks for the users to perform on the followed a simple pattern. Roughly half of the information the user was asked to find was located inside the F-shaped pattern and the other half outside the pattern. Some questions were selected to test the different menu-bars as well. The questions that were created can be seen in the pilot study (Ref pilot study).

Chapter 6

Pilot study

The goal of this pilot study is to test if the test layout will have any chance of giving any significant results. The pilot study will also test if the tasks in the study actually answers what we want to find out. The pilot study will also give a clearer indication of the limitations of these tests.

6.1 Method

The pilot study was done by showing the test people the developed sketch prototype. Using this sketch prototype the tester sat next to the user and asked them to perform the tasks written down on a piece of paper (In Chinese for the Chinese users and in English for the Western users). First the people got a minute to look around the page to get a quick feel for the layout of the site. Then a question was showed to the user and a timer was started at the same time. When the test subject found the requested image or text they indicated that they had found the information and the timer was then stopped. This was repeated until all the tasks were fulfilled.

6.2 Results

The users were asked to perform the following tasks:

English BBC Questions:

1. Click the news about ivory stabbing
2. Click on the Korean men beauty revolution
3. Click on the news about the freed samung heir
4. Click on the news about Zuma refusing to step down.
5. Click on the news that has to do with an angry sports coach.
6. Click on the long read article about the catholic priest father
7. Click on the video about cooking with strangers
8. Click on the video that has to do with Indonesia
9. Via the top menu go to the new phones site
10. Via the top menu go to US politics
11. Via the top menu go to news about the stock market

English QQ Questions:

1. Click on the following news: One hundred Hongkong staff more than half hiding in the United States and Canada
2. Click on the following news: Fishermen are no longer allowed to bring their own baits.
3. Click on the following news: Russian fighter pilots last words before blowing himself up with a grenade "For my brothers"
4. Click the following news: True beauty don't fear wrinkles
5. Click on the video with a Chimpanzee
6. Click on the video below: Premier League - Liverpool 2-2
7. Click on the skyscraper picture
8. Click on the news below: Dow plunge near 700 on Friday what triggered it?
9. Choose from the following menu items: News
10. Choose from the following menu items: Health
11. Choose from the following menu items: Sports
12. Choose from the following menu items: Digital

The BBC pilot study resulted in the following results (see fig 6.1):

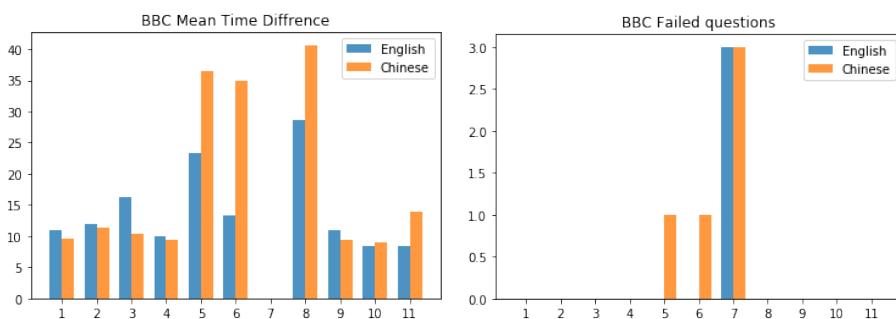


FIGURE 6.1: Results from the pilot study for the BBC inspired news prototype.

The QQ pilot study resulted in the following mean results (see fig 6.2):

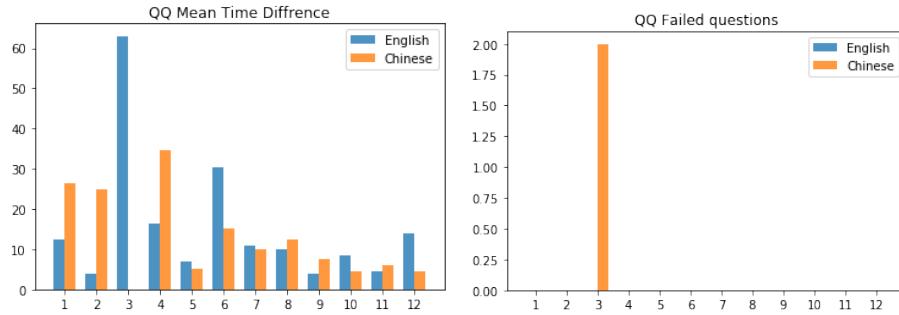


FIGURE 6.2: Results from the pilot study for the QQ inspired news prototype.

6.3 Discussion

Doing this study provided a lot of relevant information some of the main problems with the test that was identified where that some of the news where repeated on several places of the site this made some tasks irrelevant since the news could be located at several different locations. Some of the questions seemed badly translated as well. For example the question of the sports coach seemed to confuse many of the Chinese users. Also the questions regarding finding images on the qq site did not provide with any meaningful result this since QQ has very few pictures and therefore they did not check how well the user preformed in information dense sites. Another thing that was noticed during the test were how much the positioning of the questions were. The users seem to start their new search pattern from the point of the last task. This means that subjects who were asked to find new information all quickly found news closely located to the previous task. This needs to be kept in mind when designing the next set of questions, it might also be interesting to keep this in mind when analysing the results of the larger study.

We can see the measurements from the study in 6.1 and 6.2. Since the goal of the pilot study was to try out if the concept for the real study works we did not have enough participants for this data to have any statistical significance. As mention above the goal of the study was to find problems with the questions, translation and ux. According to (ref vem det nu var) we only need about 5 participants to find the majority of the user experience problems. But if we want this survey to be statistical significant when actually measuring time differences we would need 20 participants on each individual page. This would mean a total of 80 participants.

One thing that was noticed that were missing from a usability perspective is asking the user to perform actual tasks. All of the questions where focused on finding information. We would also to some extent test how the users deal with actual tasks and functions that are present on the different sites. One common task that is used on news based sites is giving feedback to the hosts and following the sites on social media.

6.4 Conclusion

Many questions will be changed to better be able to get results for the projects, also both the sites will have questions with the same structure. Both sites will have 16 tasks to perform. Four of the tasks will be about the menu-bar, four of the tasks will be functional, four of the tasks will be about finding precisely described news titles and lastly 4 of the tasks will be about finding more general described news. About half of the tasks will be in the F-shaped pattern view sight. The other half of the questions will be located to the right-hand and central side of the website. Finally the sites will be designed so the content of both sites will be as similar to each other as possible.

Some functionality will be added to the prototype such as giving feedback and also following on social media. This will be done according to standards as can be seen in 4. A menu with the option to give feedback and follow on social media will be added to the right hand side on the Chinese pages respectively on the bottom of the page for the western site.

The new questions selected for bbc are the following:

The questions selected for qq are the following:

Chapter 7

Phase 3 - Building the Interfaces

The goal for this phase is to develop the beta-test, test it, update according to feedback and deploy it. After the web test has been deployed it will be accessible. Several different technologies will be used and described in this chapter to make all of this possible.

7.1 Work process

This programming phase is the largest and most time consuming part of this thesis and there are several requirements for the tests and its results to be useful. Requirements needed for the unsupervised test to be usable are the following:

1. The test need to look the same for all users.
2. The site needs to be loaded fairly quickly and feedback from the loading process should be displayed to the user.
3. The user is not allowed to go back and redo any questions.
4. The site has to be fully loaded before the test can be started.
5. Users are not allowed to use ctrl-search.
6. What site the user will perform their test has to be controlled, this so we receive equal amount of tests for both sites.

To be able to perform the test on users all over the world without actually having to be there a web based test had to be constructed. The test was made using several different technologies and hosted on aws.

Several things had to be taken into consideration such as: slower network in china, possibility of web connection getting interrupted, measuring correct behaviours, making sure a completable devices was used for the test (mobile device would not at all test the same thing).

7.1.1 Front-End

The most important front-end technology on this website are React, Redux and Redux-Sagas. These three libraries create most of the functionality in website and they work very closely together. With React we show the user what we want him or her to see. All the users behaviours are stored in Redux (imagine a database for the browser), then depending on the updates in Redux, React appropriately up dates the information the users see. Some of the actions a user does triggers a Saga (Example: Pressing the finish button). That Saga then makes a asynchronous call to our

API and send over the data stored in Redux to our mysql database hosted on AWS in Seoul.

Our front-end consist of four different pages. Homepage, bbc, qq, sus and done. The names of these pages are taken from the material that inspired them. The bbc site is not a actual bbc site but is named so since it is inspired from the bbc website. These namings is not something the users see and therefore does not affect them it's only to make it clear what page that is currently being disused. The flow of the test is as follows: The user starts at the Homepage, depending on the last test made by someone the user will either end up at the bbc page or the qq page. Depending on the language selected by this user they will see the page in the selected language. After finishing the bbc or qq test the user will be taken to the questionnaire on the sus page. After this is completed the user will arrive at the done page which contain a simple message thanking the user for their participation and provide my contact information if they have any questions.

Homepage

The homepage is the first page the user will see. This page is responsible for gathering information needed to decide how the rest of the test will be set-up. The homepage will start by asking the user if they want to do the test in Chinese or English. The test will then proceed to give the user information about the test in their chosen language. The web page will then query the database to check which of the sites that have the least number of tests done (bbc or qq). The user will then be see a description of the test in the selected language (see [7.1](#) and [7.2](#)). After the user finishes reading the description and press "start" the site will start the test.

Welcome please answer the following questions to start.

Language: English

Gender: Female

Age: 25-34 years old

Continue

FIGURE 7.1: Image of the first page the user is greeted with.

Hi and thanks for participating in the following test.

My name is Marcus and this test is a part of my master thesis research, which explores cross-cultural website usability. The test will take roughly 5-10 minutes and the format is as follows: Once the test finished loading, you will be shown a news site. On this news site, you will be asked to find different articles and links. In some instances, you will be asked to directly click on an item based on its description. In other cases, a more general description will be given. After clicking on the link you are asked to find, please click next to proceed. If you are unable to complete a certain task, you can skip it. All solutions to the required tasks can be found on the website, so please aim to complete the task before skipping it. After the 13 tasks, you will be asked a couple of general questions on the site you used.

IMPORTANT: Please avoid using search tools since this will make your results meaningless. Also please try to do the test without interruptions since your actions and click patterns will be timed and recorded. Use a computer for this test it is not meant to be done on a phone! Avoid using the back button, you can not redo a task if you did a mistake that is fine just continue with the test.

Start

嗨，感谢您参加以下测试。

我的名字是马库斯，这个测试是我的硕士论文研究的一部分，它探索了跨文化网站的可用性。测试将花费大约5-10分钟。格式如下：一旦测试完成加载，您将看到一个新闻网站。在这个新闻网站上，您将被要求找到不同的文章和图像。在某些情况下，系统会要求您根据其描述直接点击某个项目。在其他情况下，将给出更一般的描述，点击您需要找的项目后，请点击下一步继续。如果您无法完成某个任务，则可以跳过它，所有必要任务的解决方案都可以在网站上找到，所以请在跳过之前完成任务。完成13个任务后，您会在您使用的网站上询问几个常见问题。

重要提示：请避免使用搜索工具，因为这会使结果变得毫无意义。此外，请尝试不中断地进行测试，因为您的操作和点击模式将被定时和记录。使用电脑进行这项测试并不意味着要在手机上完成！避免使用后退按钮。如果你犯了一个错误，那么你不能重做任务，只要继续测试。

开始

FIGURE 7.2: Description in Chinese and English.

BBC and QQ

Both the bbc and qq inspired sites have the same basic layout. When the user arrives at the page a pop-up will appear with a loading bar, this is to make sure the user can't start the test until all the images has been loaded from the server. When all the images have been loaded the user will be able to start the test (see A.1 and A.2 in Appendix A for full sites in English and Chinese).

All the news in this test has been translated so they have both the English and Chinese versions available depending on what language the user selected from the Homepage. Once the user starts the test a timer starts, the user is not able to see this timer. Each click the user does is recorded, the time it takes to complete the task and if the user gets the correct task or not. The number of questions the user have left to complete is also shown. The questions and what the user has most recently clicked is shown in a bar at the bottom of the screen (see 7.3). Once the user finishes the test their answers are sent to the database via the API and they are redirected to the SUS site.

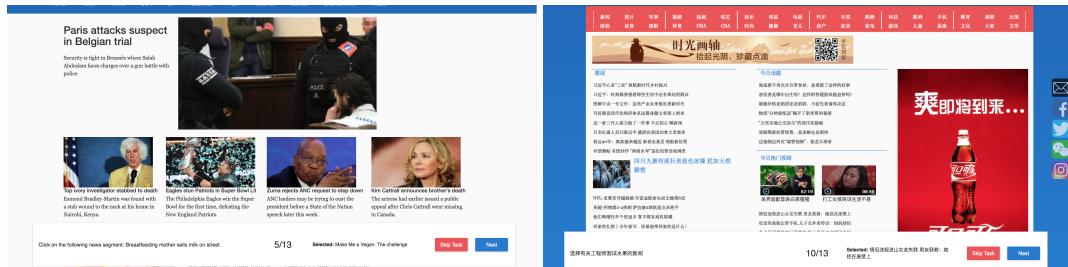


FIGURE 7.3: A example of what it can look like for a user when testing bbc and qq inspired websites.

Questionnaire

This site consists of the before mentioned predefined questions designed to get a understanding for what the user feel about the website, test and design (See fig 7.4). Once the user has answered all the questions and pressed finish their answers will be sent to the database via the api. The user will then be rerouted to the done site where they will see a message thanking them for participating in the test (See fig A.3 in Appendix A).

The screenshot displays a series of ten survey questions, each with a five-point Likert scale from 'Strongly Disagree' to 'Strongly Agree'. The questions are as follows:

- I liked the design of the site / 我喜欢网站的设计
- The design of this site was similar to other news sites / 这个网站的设计与其他新闻网站类似
- I think that I would like to use this site frequently / 我认为我想经常使用这个网站
- I thought the site was easy to use / 我认为该网站很容易使用
- The design of this site was unusual to me / 这个网站的设计对我来说是不寻常的
- I thought there was too much inconsistency in this site / 我认为这个网站有太多不一致的地方
- I felt very confident using the site / 我对使用该网站非常有信心
- I thought the material I was looking for was easy to find / 我认为我寻找的材料很容易找到
- I found the site very cumbersome to use / 我发现该网站使用起来非常麻烦
- I thought that the amount of information on this site was too sparse / 我认为这个网站上的信息量太稀少
- I felt overwhelmed using this site / 我感到不知所措使用这个网站

Each question includes radio buttons for responses 1 through 5, with 'Strongly Disagree' at 1 and 'Strongly Agree' at 5. A 'Finish Test' button is located at the bottom right.

FIGURE 7.4: The Questionnaire site.

7.1.2 Database

The database consists of five different tables: Main, Actions, Questions, Sus and QuestionTexts. (See fig: 7.5 for figure of the full database)

The Main mainly is used to identify the user where each user has a unique row. The table has four columns Id, Site, Language and Age. The id in the main table is the main id to identify a user. This same id can be found for each user in all other tables except QuestionTexts.

The Questions table contain all the answers from the users, the table has seven columns: Id, MainId, QuestionId, Correct, StartTime, EndTime, Correct, TotalTime.

MainId and QuestionId is foreign keys referencing the ids in the Main and QuestionsTexts tables. Each row in the Question table contain a answer from a user.

The Action table contain all the click-actions a user did per question. The table contain seven columns: Id, QuestionsId, PosX, PosY, ScreenWidth, ScreenHeight, RelativeTime. Where QuestionsId is a foregin key referencing the Id in Questions. Each row in the Action table contain a action made by the user.

The Sus table simply contains all the users answers to the Sus questions. The QuestionTexts table contain all the questions the user are asked in the test

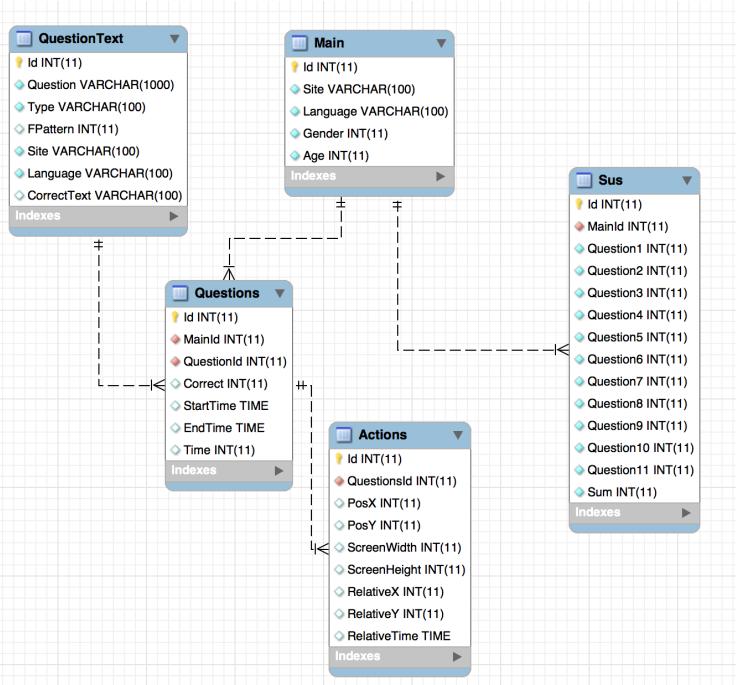


FIGURE 7.5: A EER schema of the mysql database. Yellow icons show the tables primary key, red the foreign keys and blue/white the attributes.

7.1.3 Api

A express Api was set up to handle communication between the front-end and database. The Api takes the information that is sent from the users front-end code and transform it into a format that works for the sql database. The Api then quires the database and insert the new data into the database tables.

7.1.4 Hosting AWS

To host this app on AWS we used the feature called "Elastic Beanstalk" also called eb. Eb allowed us to easily launch a application that automatically set up a EC2 instance, auto scaling, load-balancing, RDS. The services was set up in Seoul. This to decrease the loading time for China as much as possible. Since Seoul also have very good network connections to the rest of the world this does not increase the loading time

in Europe and the US that much. A lot of time had to be spent changing parts of the code so it would run on AWS servers.

7.1.5 Beta-tests

Once a working version of the test was created. A beta test was made to find possible bugs in code and improve user-experience in terms of font-size and design of the testing parts. The beta-tests were made in the form of letting users try to finish the test while it was supervised. Notes were taken about possible misunderstandings, bugs and improvements that could be made to the test. Some examples of notes that were taken during the beta tests can be seen below.

1. Make selected bigger so the user can easier see what they have done.
2. Should we log scrolling?
3. Some correct question gets logged as incorrect in database in-spite of being correct.
4. Waiting for sus site is very slow. Is it waiting for the database?
5. Left top-side of qq site is too small, because of commercial?
6. Wrong spelling in some news.
7. Skip task not working.
8. Some correct question gets loged as incorrect in database in-spite of being correct.
9. Change so the size of the page is constant no matter the screen size. (Looked very bad on a big screen right now.)
10. Start time not working correctly on some questions.
11. Remove video from questions, does not give any relevant information.

The beta was tested on about ten people where six did the English version of the test and four the Chinese version. After about eight tests almost no new information about usability problems and bugs were found so it was concluded to finish the beta testing and launch the test.

7.1.6 Launch

The launch of the test went fairly quickly with AWS. During the launch the performance of the site was monitored closely and we could thereby see that loading times increased when several people used the site at the same time and not all results were logged in the database. It was quickly concluded this was because a problem with the database connection from the api and within the matter of minutes this bug was fixed and the site was updated. It can be estimated that about 4-5 test results were lost due to this mistake.

7.1.7 How the test was conducted

The test was conducted by the users completely unsupervised. A link was sent to all the users asking them to click the link and perform the test. The users were asked to only perform the test once and on a computer. The rest of the information about the test was described in the first page of the test (see fig: 7.2)

Chapter 8

Phase 4 - Analyzing Data

The goal with this phase is to use the language python combined with mysql quires to get the data from the database. Clean the data and then analyse the results and test it for statistical significance.

8.1 Work process

The main technologies used for this phase was Python using several different libraries and mysql to querie the database in Seul, South Korea.

data gathering process: To gather the data from the MySQL database several queries were made with conditions that specified what specific data to be gathered. This data then loaded into a data frame in python. Eight different data frames were created. Four containing information about the user's actions on each question these were the following: English speaking BBC users, Chinese BBC users, English QQ and Chinese QQ users. Four other data frames with the same constraints were created but instead of containing the user's responses and times it took to answer the different questions, these contained the users questionnaire answers.

Once the data was gathered is was scanned for irregularities these were often either removed or replaced by a mean value as is commonly done in machine learning and other data mining/data analysis fields. The most obvious errors were the ones where the data had the wrong shape because of a bug. These were simply fixed or removed one example of this is one user who had twice the amount of answers to the questions with the exact same information. In that case, one of each row containing a duplicate was simply removed. Another type of faulty data where the user's complete input was removed was the cases where a user had skipped through all his/her questions. Resulting in all incorrect answers with times spent on each question close to zero, these users were simply removed another type that was removed were users who did the test from a mobile device. From this type of data cleaning, 16 users were removed (total users was calculated after these had been removed). Another type of data that needed to be cleaned where some questions where the users simply had an extremely high mean time of a question that was out of proportion to the rest of the mean times. This seemed to happen randomly and is likely due to a user leaving the test to do something else and then return. In these cases, a cap of 400 seconds where set and any question mean-time that was higher than that was removed. The 400-second cap was used because the largest normal average was around 100-200 seconds after that we only got a couple of odd outliers with times from 500 seconds and up. There were some smaller oddities that had to be handled as well but we will not go into these in detail in this report. The data was removed by two different ways, if a user had to be removed it was simply unselected in the

SQL query to the database. If on the other hand a value had to be modified or a single question result dropped this was done in the python script on the data frame.

8.2 Analysis

To analyse whether the differences in mean correct results from the BBC site and the QQ site were statistically significant, I first calculated the average proportion of correct responses from each respective site per question. Specifically, I used Python to find the average correct responses for each group as well as to plot these results on a bar graph (see figure xx). Additionally, for both sites, I calculated the average amount of time it took per user to submit their answer given that the users' answer was correct. The number of seconds it took users per question, on average, to find the accurate answers were then plotted as bar graph using Python (see figure xx). Finally, to examine whether the differences in mean time and proportion of accurate answers for English versus Chinese users on the BBC and QQ sites were statistically significant, I performed an independent samples t-test analysis on the mean values using Python.

8.3 Questionnaire

At the end of the study, users were requested to complete a set of 11 self-reported questionnaire questions gauging the likeability of the site design (e.g. "I liked the design of the site", "The design of this site was unusual to me", etc.) and perceived navigability of the site (e.g., "I thought this site was easy to use", "I felt very confident in using the site", etc.). Users were asked to respond to the questions on a 1 to 5 Likert scale, where 1 reflected strongly disagree and 5 represented strongly agree. However, the directionality of the responses - if they reflected positive or negative perceptions of the sites - depended on the weight of the questions (i.e., "I liked the design of the site" reflects a positively weighted question while "I felt overwhelmed using this site" reflects a negative weight). Given that the directionality of the questions was not uniform, I first normalized the data prior to calculating the overall likeability of the BBC and QQ sites. Additionally, I dropped responses from question two as this question simply asks users whether they believe the news site was similar to other news sites, as this does not measure likeability or navigability of the site, it was excluded from this analysis.

Chapter 9

Results

Each test was only done once per user meaning that all the users only saw the QQ site or the BBC site in their language. The questionnaire contained the same questions no matter which site the user performed their test on.

9.1 BBC English

For the BBC English website, there were a total of 29 participants. Most respondents were in the age ranges of 18-24 and 25-35 with few outliers. All users who saw the site were self-proclaimed English speakers and were asked to navigate the BBC website using English, these users only saw the BBC English site and not the BBC Chinese site. On average, native English speakers using the BBC site, correctly responded to task queries 88% of the time (see fig 9.1). Additionally, the mean time per question used to answer questions accurately was 21.5 seconds. To complete the tests, users took an average of 5.5 minutes, where the mean time spent per question was 25.3 seconds (see fig 9.1). For questions that fell within the F-shaped pattern, participants used roughly 19.1 seconds to answer a question. Conversely, users took 31.8 seconds to answer questions outside the F-shaped pattern. The difference between how long it took English users to find answers within the F-shaped pattern (19.1 seconds) and outside the F-shaped pattern (31.8 seconds) proved to be statistically significant ($t = 2.8479$, $p < 0.05$), indicating that English speakers were quicker when navigating the BBC site within the F-pattern compared to outside.

The Questionnaire results can be seen in 9.7. The overall calculated result for the BBC English site based on the Questionnaire is 4.88. Where 0 is the lowest score and 10 is the highest.

9.2 BBC Chinese

For the BBC Chinese website, there were a total of 21. Again, most respondents were in the age ranges of 18-24 and 25-35 with few outliers, indicating a rather homogeneous response group. All users who saw the site were self-proclaimed Chinese speakers and were asked to navigate the BBC website using Chinese, these users only saw the BBC Chinese site and not the BBC English site. On average, native Chinese speakers using the BBC site, correctly responded to task queries 74% of the time (see fig 9.1). Additionally, the mean time per question used to answer questions accurately was 24 seconds. To complete the tests, users required roughly of 5.9 minutes, where the mean time spent per question was 27.3 seconds (see fig 9.1). For questions that fell within the F-shaped pattern, Chinese participants used roughly 24.4 seconds to answer a question. Conversely, Chinese users took 28.6 seconds to

answer questions outside the F-shaped pattern. The difference between how long it took Chinese users to find answers within the F-shaped pattern (24.4 seconds) and outside the F-shaped pattern (28.6 seconds) proved to be statistically significant ($t = 5.3301$, $p < 0.05$), indicating that Chinese speakers were, surprisingly, slightly quicker when navigating the BBC site within the F-pattern compared to outside.

The Questionnaire results can be seen in 9.7. The overall calculated result for the BBC English site based on the Questionnaire is 5.51. Where 0 is the lowest score and 10 is the highest.

9.3 Differences between Chinese and English BBC users

When comparing the results of Chinese users and English users for BBC, English users, who answered 88% of questions accurately, tended to get questions correct more often. In comparison, Chinese users only answered 74% of questions correctly on the BBC site. The difference in the number of correct responses were found to be not statistically significant at the $p=0.05$ level of significance ($t = 1.8893$, $p>0.05$). Interestingly, there appeared to be no difference between the proportion of questioned answered correctly between the Chinese and English BBC users. Additionally, although Chinese users spent more time on average to respond to questions, taking an average of 5.9 minutes whereas English users only took 5.5 minutes, the two groups were not statistically different in their mean response time rates ($t = -0.49$, $p > 0.05$). These results indicate that there is no major difference, beyond random chance, that Chinese users or English differ from each other in terms of time spent on answering questions and the tendency to answer questions correctly.

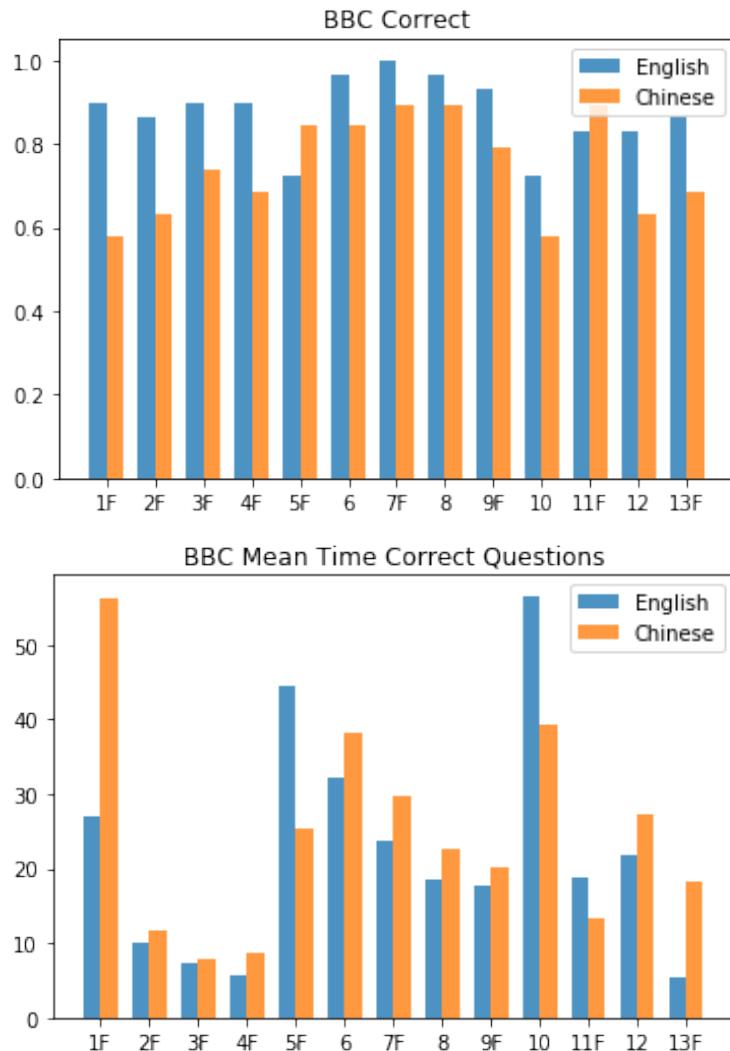


FIGURE 9.1: The Correct and Mean time per correct question results from the bbc page.

9.4 QQ English

For the QQ English website, there were a total of 30 participants. All respondents were in the age ranges of 18-24 and 25-35. All users who saw the site were self-proclaimed English speakers and were asked to navigate the QQ website using English, these users only saw the QQ English site and not the QQ Chinese site. On average, native English speakers using the QQ site correctly responded to task queries 87% of the time (see fig 9.2). Additionally, the mean time per question used to answer questions accurately was 17.7 seconds (see fig 9.2). To complete the tests, users took an average of 4.3 minutes, where the mean time spent per question was 19.3 seconds. For questions that fell within the F-shaped pattern, participants used roughly 17.6 seconds to answer a question. Conversely, users took 20.7 seconds to answer questions outside the F-shaped pattern. The difference between how long it took English users to find answers within the F-shaped pattern (17.6 seconds) and outside the F-shaped pattern (20.7 seconds) proved to be statistically significant ($t = -3.1606$, $p < 0.05$), indicating that English speakers were quicker when navigating the QQ site

within the F pattern compared to outside.

The Questionnaire results can be seen in 9.7. The overall calculated result for the BBC English site based on the Questionnaire is 2.62. Where 0 is the lowest score and 10 is the highest.

9.5 QQ Chinese

For the QQ Chinese website, there were a total of 21 participants. Again, all respondents were in the age ranges of 18-24 and 25-35, indicating a rather homogenous response group. All users who saw the site were self-proclaimed Chinese speakers and were asked to navigate the QQ website using Chinese, these users only saw the QQ Chinese site and not the QQ English site. On average, native Chinese speakers using the QQ site correctly responded to task queries 85% of the time (see fig 9.2). Additionally, the mean time per question used to answer questions accurately was 18.5 seconds (see fig 9.2). To complete the tests, users required roughly of 4.9 minutes, where the mean time spent per question was 22.5 seconds. For questions that fell within the F-shaped pattern, Chinese participants used roughly 22.6 seconds to answer a question. Conversely, Chinese users took 22.4 seconds to answer questions outside the F-shaped pattern. The difference between how long it took Chinese users to find answers within the F-shaped pattern (22.6 seconds) and outside the F-shaped pattern (22.4 seconds) proved to be statistically significant ($t = -5.907, p < 0.05$), indicating that Chinese speakers were, slower when navigating the QQ site within the F pattern compared to outside.

The Questionnaire results can be seen in 9.7. The overall calculated result for the BBC English site based on the Questionnaire is 4.46. Where 0 is the lowest score and 10 is the highest.

9.6 Differences between Chinese and English QQ users

When comparing the results of English users and Chinese users for QQ, English users, who answered 87% of questions accurately, tended to once again get questions correct more often. Again, by comparison, Chinese users, who got 85% of questions correct, answered fewer questions accurately on the QQ site. The difference in the number of correct responses were not statistically significant at the $p=0.05$ level of significance ($t = 0.42, p>0.05$). Interestingly, there appeared to be no difference between the proportion of questioned answered correctly between the Chinese and English QQ users, just as for the BBC users. Additionally, although Chinese users spent more time on average to respond to questions, taking an average of 4.9 minutes whereas English users only took 4.3 minutes, the two groups were not statistically different in their mean response time rates ($t = -0.49, p > 0.05$), once again, the results are similar to BBC. These findings suggest there is no major difference, beyond random chance, that Chinese users or English differ from each other in terms of time spent on answering questions and the tendency to answer questions correctly.

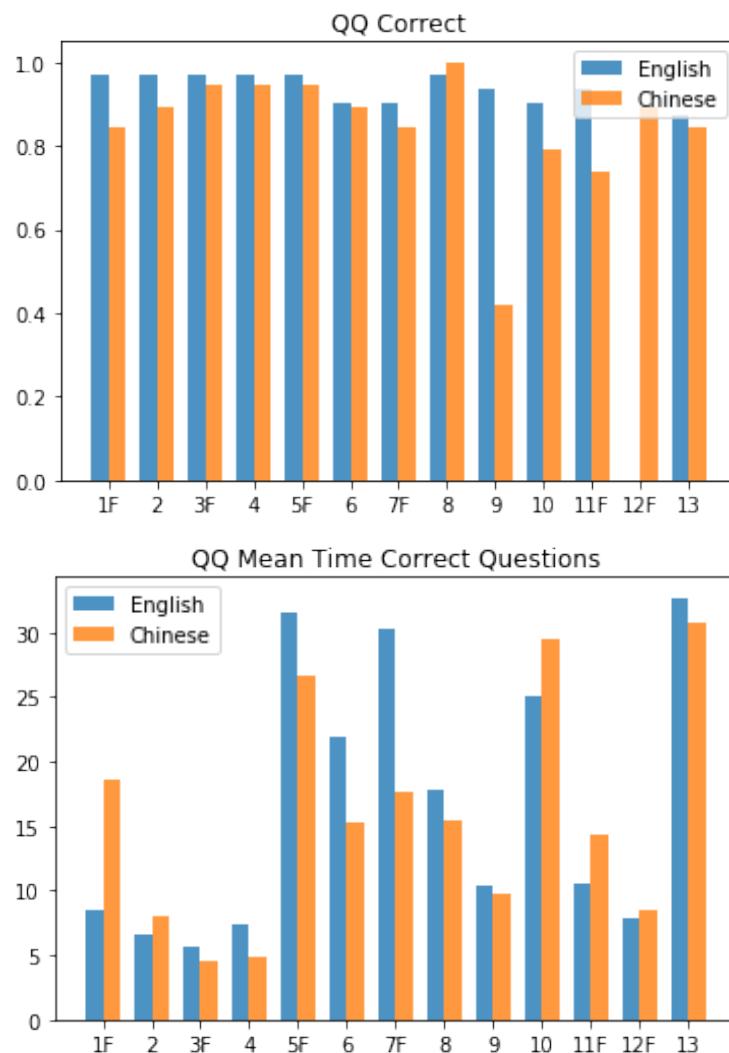


FIGURE 9.2: The Correct and Mean time per correct question results from the qq page.

9.7 Questionnaire Results

Question	QQ Chinese	QQ English	BBC Chinese	BBC English
1: I liked the design of the site	2,52	1,6	2,95	2,34
2: The design of this site was similar too other news sites	3,9	1,97	3,24	3,45
3: I think that I would like to use this site frequently	2,19	1,33	2,33	1,97
4: I thought the site was easy to use	2,48	1,6	3,14	2,72
5: The design of this site was unusual to me	2,71	3,73	3,2,52	2,41
6: I thought there was too much inconsistency in this site	3,1	3,67	2,71	2,89
7: I felt very confident using the site	2,52	2,1	2,9	2,76
8: I thought the material i was looking for was easy to find	2,38	1,73	2,9	2,48
9: I found the site very cumbersome to use	3,48	3,93	3,05	3,07
10: I thought that the amount of information on this site was too sparse	2,43	2,33	2,9	2,69
11: I felt overwhelmed using this site	2,81	4,23	2,71	2,76

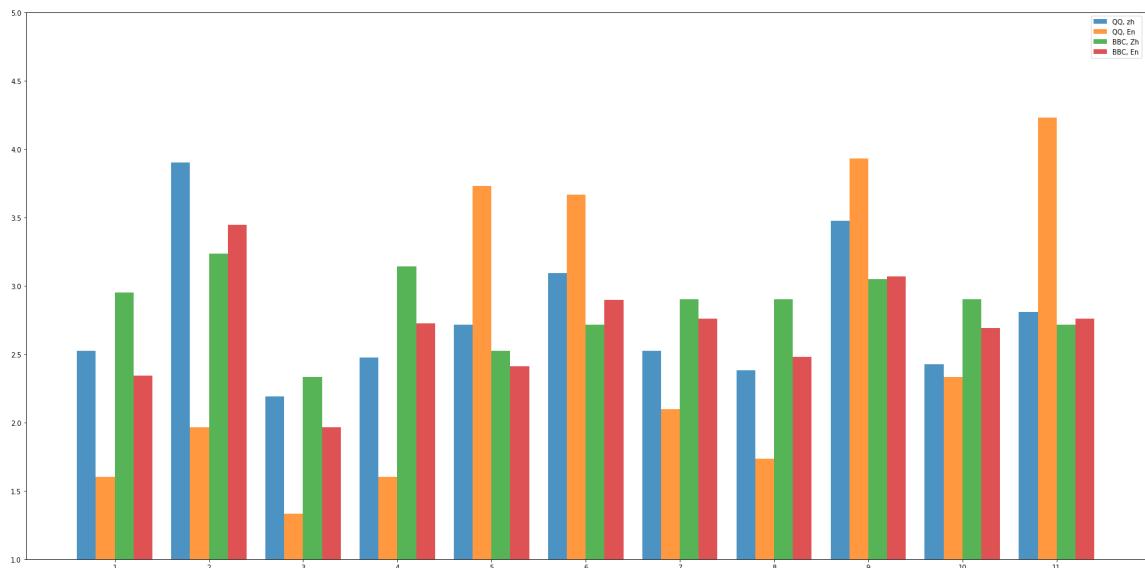


FIGURE 9.3: The results from the Questionnaire

Chapter 10

Discussion

In this research, I aimed to investigate 1) how differences in interface design may be due to differences in information processing styles or simply trends, 2) how do different processing styles in Western (analytical) versus Chinese (holistic) users affect performance on different interfaces, and 3) if one global interface be created, or if web designers focus on creating separate user interfaces for different cultures. Concerning the question of how interface design may be due to different information processing styles or trends, my research suggests that a mix of both factors have contributed to the different web designs from China and its western counterparts. Specifically, Chinese users preferred the western based design over the Chinese one, according to self-reports [report stats XXX]. Interestingly, this suggests that perhaps the reason why Chinese news sites tend to be more information dense may simply be due to a trend in web design rather than fundamental perception differences. For instance, Chinese users performed as well as English users navigating both the BBC site ($t = -0.49, p > 0.05$) and QQ site ($t = 0.42, p > 0.05$) indicating that there is no statistical difference between the two groups.

Regarding the question of how information processing styles western (analytical) versus eastern (holistic) would affect user performance on different interfaces, my research suggests that information processing style does significantly influence performance. Unsurprisingly, my research indicated that English, or analytical, users who used the BBC site were much quicker at finding important objects within the F-shaped pattern compared to outside of it ($t = 2.8479, p < 0.05$), taking an outstanding 12 seconds longer outside the F-shaped pattern. This is congruent with the existing literature on UX, which tends to be western-centric, as eye tracking studies suggest western users tend to scan through a website using an F-shaped pattern. However, it was surprising to find that Chinese users were actually also able to answer questions more quickly when questions laid within the F-shaped pattern versus outside ($t = 5.3301, p < 0.05$). It appears that for both Chinese and English users of the BBC site, objects that fell within the F-pattern were easier to find than those outside. However, it appears that the effect size was significantly larger for English users, with a 12 second difference inside and outside the F-pattern, than for Chinese users, who only had a 4 second difference. Overall, it appears that a western based design seems easily navigable for both eastern and western users and the f-shape appears to increase performance for both groups. The main difference, however, is that English users were significantly better within the F-pattern (taking 12 seconds less) than outside, whereas Chinese users only saw a slight improvement (taking only 4 seconds less within the F-shape).

On the effects of information processing and website navigability, my findings suggest that English speakers using the QQ site were quicker inside the F-shaped pattern than outside of the pattern ($t = -3.1606$, $p < 0.05$), with a 3 second difference. For Chinese users on the QQ site, conversely, they were very marginally slower inside the F-pattern (using 22.6 seconds) than outside (22.4 seconds) the pattern ($t = -5.907$, $p < 0.05$), suggesting that Chinese users using the QQ site did not rely on using the F-shaped pattern as much as western users. Overall, it appears that even though both Chinese and English users were able to navigate the QQ site with the same speed ($t = -0.49$, $p > 0.05$), English users were still slightly more dependent on the F-pattern, like the findings from the BBC site. Interestingly, however, according the questionnaire self-reports, it appeared that Chinese users found both the QQ and BBC websites to be equally likeable 4.46 (QQ) vs 5.51 (BBC) and did not seem to be overwhelmed by the information density of the QQ site, average of 2.8 from 1 to 5 on the question "I felt overwhelmed using this site". For the English users, however, self-reports indicated that they were highly uncomfortable with the layout of the page, finding the information overwhelming. They had a average of 4.23 from 1 to 5 on the question "I felt overwhelmed using this site". They also tended to give higher ratings for BBC than QQ 4.88 (BBC) vs 2.62 (QQ).

Taken together, these results suggest that information processing does influence how users interact with different interfaces. It appears that while western users can use a Chinese designed site, they find the experience to be extremely unpleasant 2.62 (QQ). Conversely, my findings indicate that Chinese users find both sites equally likeable 4.46 (QQ) vs 5.51 (BBC). Furthermore, English users were more reliant on the F-shaped pattern than Chinese users were, which is in-line with existing research on user perception. Concerning the question of whether or not one global interface should be created or if web designs should be tailored to different cultures, my research suggests that one global interface can be deployed to maximize efficiency. Chinese users found BBC to be just as likeable as QQ, whereas English users disliked the information density of QQ, suggesting that a western designed site with sleek information layouts appear to be generally liked by users. Accordingly, it appears that web designers can focus on creating one global interface rather than tailoring websites to different cultures.

Chapter 11

Conclusion

This research study was conducted virtually, without the researcher interacting directly with participants or observing their behaviors. As UX research generally explores how users interact with an interface, researchers have traditionally examined how people navigate an interface through direct observation. As such, my decision to run this study virtually lends itself to certain disadvantages in UX research. For instance, I am not able to talk to my participants in-depth to record their individual navigation patterns and perceptions of the website. However, although running my study virtually caused me to lose depth from each individual, running the study at scale offers significant advantages in understanding user preferences as well. For instance, I can run the study at a larger scale, meaning that I was able to get more participants to take the test, which was necessary for me to obtain generalizable statistical results. Further, being able to conduct the study with a larger sample size is better suited toward my specific research questions as I am exploring how people from different cultural background may perceive a given interface differently. Considering my research question, it would be difficult to draw generalizable conclusions with a small sample size of only 5-10 participants, as is common in UX research. In this way, I can examine more of the general trends that emerge when users from China or the US interact with the interface.

In this research, I was able to find interesting explanations to my research questions. Specifically, I found first that 1) differences in interface design is not particularly culturally dependent or trend dependent as web design is a mixture of both factors. For instance, my analysis indicates that although Chinese users reported liking the Chinese designed site more than westerners, both groups still performed equally well on both the western and Chinese sites. Second, I find that cultural differences do influence user performance on different websites. Western people performed significantly better when answering questions that fell within the F-shaped pattern with a much larger difference in mean times than Chinese. Although both groups were quicker with finding solutions within the F-shaped pattern on the BBC site, western users were a striking 12 seconds faster within the F-pattern, compared to Chinese users who were only 4 seconds faster on BBC. Additionally, on the QQ site, Chinese users did not follow the f-shaped pattern, suggesting they employ holistic information processing. Western users on the QQ site, however, were still heavily dependent on the F-pattern and performed better within it than outside. Finally, my research actually implies that a global interface can be deployed for people of different cultures if it follows a more clean and sleek design. While on the self-report questions, westerners indicated that they found the information density overwhelming on the QQ site, Chinese users reacted positively toward both the western and eastern based websites. Further, there was no statistically significant difference regarding overall performance between the two groups, suggesting a global interface can be deployed

so long as it follows a more sleek design with lower information density.

11.1 Limitations

Although the findings of my research are compelling and the overall project design was deployed successfully, there were still factors that limited the overall strength of the findings. First, as this research had to be done under time constraint, there simply was not enough time to develop a perfect site that captured all the variables I was interested in exploring. Given more time and resources for this project, I would have been interested to supplement my findings with eye-tracking in order to determine if the heat maps would reveal different patterns for Chinese and western users. Further, I was limited in receiving feedback from users considering the test was done virtually. Although doing this testing virtually provided me with more data and generalizable data, I was not able to get the same depth of reflection from users. If the test were conducted in-person, I could have had an in-depth interview with participants to understand their deeper perceptions concerning these different websites. However, because I am personally based in Sweden whereas many of my participants were from China, I was unable to conduct one-on-one in-depth interviews. Finally, the majority of people who participated in my research were college students or individuals in the 20-30 age range. As I did not have many participants who were greatly older or younger than this range, it may be difficult to generalize the findings to older populations.

11.2 Future work

In terms of future work on this subject, I believe that adding in-depth interviews would greatly contribute to the overall findings of this paper. Asking users directly about their thoughts on the QQ and BBC sites would give richer insight into how people from different cultures feel about the the web designs. Currently, I am basing user perception differences on self-report questionnaire questions and click-time measurements. In the future, I would like to explore users personal opinions about the subject to add more depth. In addition, my study only used two web pages, BBC and QQ, to study underlying perception differences for people of different cultures. Future research should run the same experiment with more variation and an increased number of web pages to see how well my results hold up if the content and web designs are varied. Furthermore, my studies only had roughly 50 users per site. Although this is a good amount of participants for a small-scale study, I believe that adding more participants would increase the validity of the tests. As such, I believe this is a good domain for future researchers to explore or for me to extend in my own projects.

Appendix A

Appendix A



FIGURE A.1: The full view of QQ site in english and chinese

News

巴黎攻击怀疑在比利时审判

当地时间1月10日，涉嫌策划巴黎连环恐怖袭击的萨拉赫·阿卜杜勒·马吉德在比利时首都布鲁塞尔出庭受审。他被控犯有组织、准备和实施暴力恐怖行为罪，以及谋杀罪。

美国参议院投票通过《2017年国防授权法案》，该法案将为美国军方提供约7000亿美元的预算，以应对朝鲜和俄罗斯的军事威胁。

1/13 Selected: 起诉谷歌母公司涉嫌反垄断案
Next

新闻菜单选择: 电视
新闻
Waze v Uber: 谷歌做了什么? 英国诉讼出关键同盟, 10号呈报

必看
作为承办人的生活... 17岁
Korean men are starting a beauty revolution
Meet the autistic models storming the catwalk
3 菲利普斯试图解释
全文
从断脖子到牛津的罗兹奖学金
为什么我们没有飞行是缺陷?
2 遇见自闭症模特们登台表演
4 作为承办人的生活... 17岁
3 菲利普斯试图解释
长闻
大部分读
在世界各地
寻求像疾病一样治愈衰老
运动
新闻
在这里找找
行

News

Paris attacks suspect in Belgian trial

Security is tight in Brussels where Salah Abdeslam faces charges over a gun battle with police.

Two men are charged with the killing of Belgian police officer Xavier Jooris in October 2014. The suspect, known as "The Spider," was arrested in Paris last year.

Kim Cattrall announces brother's death

The actress had earlier issued a public appeal after Chris Cattrall died following a long illness.

Click on the following news segment: Breastfeeding mother sets into street

5/13 Selected: Make Me a Vegan: The challenge
Skip Topic Next

Korea jail gives birth

IOC disappointed Russian bars were fined
Crashed Amtrak train 'on the run'
Uber v Uber: Who stole what?
UK will leave customs union, No 10 insists

Must See

Life as an undertaker... aged 17
Meet the autistic models storming the catwalk
Why have we got a cold shortage?
Prostate cancer symptoms explained
Full Story

From a broken neck to a Rhodes Scholarship at Oxford
Timberlake touches down at Super Bowl
Five things about Paris attacks suspect Chris Cattrall
The Iron Man (5)
The last reviving street cleaner in Nigeria
The mystery of the strongest billionnaire
Star general Oshawa

Long Read

1 Kim Cattrall announces brother's death
2 Paris attacks suspect in Belgian trial
3 Top ivory investigator stabbed to death
4 Snow delayed travel as temperatures fall
5 UK will leave customs union, No 10 insists
6 The unwanted holiday homes owners can't give away
7 Reality TV star Kylie Jenner gives birth
8 Timberlake touches down at Super Bowl

World

Shore's 5' giant can cost \$7,200
China's last green frog
The quest to cure ageing like a disease
The insect that painted Easter Island
The chameleon size of an ant
The hidden message in social media
How food has changed the way we shop

Sport

Angry Klopp fumes after 'Kore' penalty
Bay 15-16 England
Monday's football catch-up
State of Sport
NFL player killed by a suspected drunk driver
Timelies: Ford and Fender take permanent to new heights
Garth Crooks' team of the week

Newsbeat

Don't be afraid: 17-year-old underwriter
Make Me a Vegan: The challenge
NFL star vehicles unlikely with Meek Mill
Rapper plays the morning in court

Find us here

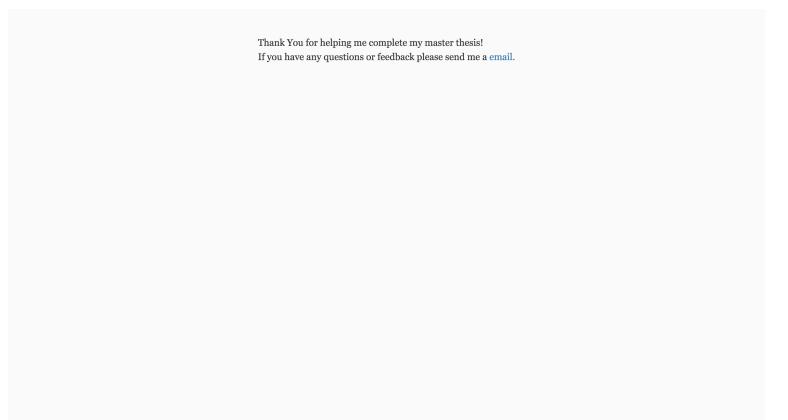


FIGURE A.3: Last site the user is shown when they have finished the test.

Bibliography

- [1] *Amazon Web Services (AWS) - Cloud Computing Services*. URL: <https://aws.amazon.com/>.
- [2] Barry C. Arnold. *Advances in mathematical and statistical modeling*. Birkhauser, 2008.
- [3] *AWS Documentation*. URL: <https://aws.amazon.com/documentation/>.
- [4] BBC. URL: <http://www.bbc.com/news>.
- [5] John Brooke et al. "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.
- [6] Dianne Cyr, Milena Head, and Hector Larios. "Colour appeal in website design within and across cultures: A multi-method evaluation". In: *International Journal of Human-Computer Studies* 68.1 (2010), pp. 1–21. ISSN: 1071-5819. DOI: <https://doi.org/10.1016/j.ijhcs.2009.08.005>. URL: <http://www.sciencedirect.com/science/article/pii/S1071581909001116>.
- [7] Ying Dong and Kun-Pyo Lee. "A cross-cultural comparative study of users' perceptions of a webpage: With a focus on the cognitive styles of Chinese, Koreans and Americans". In: *International Journal of Design* 2.2 (2008).
- [8] Li-Jun Ji, Kaiping Peng, and Richard E Nisbett. "Culture, control, and perception of relationships in the environment." In: *Journal of personality and social psychology* 78.5 (2000), p. 943.
- [9] Yuri Miyamoto, Richard E Nisbett, and Takahiko Masuda. "Culture and the physical environment: Holistic versus analytic perceptual affordances". In: *Psychological Science* 17.2 (2006), pp. 113–119.
- [10] Alokparna Basu Monga and Deborah Roedder John. "Cultural differences in brand extension evaluation: The influence of analytic versus holistic thinking". In: *Journal of Consumer Research* 33.4 (2006), pp. 529–536.
- [11] Jakob Nielsen. *Usability Metrics*. URL: <https://www.nngroup.com/articles/usability-metrics/>.
- [12] Richard E. Nisbett and Yuri Miyamoto. "The influence of culture: holistic versus analytic perception". In: *Trends in Cognitive Sciences* 9.10 (2005), pp. 467–473. ISSN: 1364-6613. DOI: <https://doi.org/10.1016/j.tics.2005.08.004>. URL: <http://www.sciencedirect.com/science/article/pii/S1364661305002305>.
- [13] Donald A. Norman. *The Design of Everyday Things*. New York, NY, USA: Basic Books, Inc., 2002. ISBN: 9780465067107.
- [14] Kara Pernice. *F-Shaped Pattern of Reading on the Web: Misunderstood, But Still Relevant*. Nielsen Norman Group, 2017.
- [15] Kara Pernice, Kathryn Whitenton, and Jakob Nielsen. *How People Read on the Web: The Eyetracking Evidence*. Nielsen Norman Group, 2014.

- [16] QQ. URL: <http://www.qq.com/>.
- [17] React, a javascript library for building user interfaces. URL: <https://reactjs.org/>.
- [18] Redux. URL: <https://redux.js.org/>.
- [19] Redux-Saga. URL: <https://redux-saga.js.org/>.
- [20] Fei Shen. "Great firewall of China". In: *Encyclopedia of social media and politics* 22 (2014), pp. 599–602.
- [21] *Top Sites in China* The sites in the top sites lists are ordered by their 1 month Alexa traffic rank. URL: <https://www.alexa.com/topsites/countries/CN>.
- [22] *Top sites ranking for all categories in China*. URL: <https://www.similarweb.com/top-websites/china>.
- [23] Tom Tullis and Bill Albert. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Elsevier/Morgan Kaufmann, 2011.