

VARIABLES AND THEIR VALIDITY

5.1 Introduction

In contrast to the theoretical evidence in the literature, the experiment (discussed in the previous chapter) did not provide empirical evidence that a relationship exists between subjective culture and usability. A number of variables that could have influenced the performance of the test subjects participating in the experiment were identified. As a result, we revised our original research problem to focus on the development of a more comprehensive model of usability that could form the basis of further research efforts in this area.

In order to develop the model of usability, it was necessary to establish the validity of the variables that were proposed to influence the experiment, as well as identify any additional variables that could influence usability. As discussed in section 1.4, a theory-building research design was followed for this research problem. A literature review served as the primary research method.

In general, very little empirical support was found in the literature for the validity of most of the variables discussed in this chapter. However, one way to establish the validity from a theoretical perspective is to determine whether or not each variable can be related to the variables that are known to influence usability, for example, context of use variables and performance determinants. Thus, the literature review is supported by the use of inductive and analogical arguments.

In establishing their validity, therefore, we identified three types of variables:

- (1) Variables that are valid, meaning that empirical, irrefutable evidence was found for the validity of the variables.
- (2) Variables that should be considered, meaning that there was no empirical evidence available, but that sufficient theoretical evidence was established to suggest that researchers should be sensitized to the potential influence of the variable on usability.
- (3) Variables that should be rejected, meaning that there is insufficient empirical and theoretical evidence that the variable influences usability.

The results of the literature investigation are presented in this chapter. Each variable or category is discussed in terms of:

- an explanation,

- why it was thought to be valid;
- how it could have affected the results of the experiment;
- what the implications are on user interface design in general;
- support available for the validity; and
- whether or not it should be accepted as valid.

The validity of these variables impacts on the way in which experiments of this nature are conducted. This is discussed in section 5.8. We conclude the chapter by discussing the influence of the valid variables on the validity of the results reported by some of the studies used from the literature to support this research (section 5.9).

5.2 Variables relating to Subjective Culture

Four variables relating to subjective culture that need to be validated for inclusion into the conceptual model of usability were identified in Chapter 4. These are cultural dimension strengths, cultural dimension interplays, the relative impact of cultural dimensions on usability, and other subjective cultural dimensions. Additional theoretical support for the validity of three of these variables was established, whilst empirical support for only one was found in the literature that we investigated. Consequently we conclude that these dimensions should be considered for inclusion into the usability model. We will now discuss each of these four variables and the evidence provided in the literature in more detail.

5.2.1 Cultural Dimension Strengths

This variable considers the possibility that cultural dimensions are present in varying strengths in different users. Hofstede's [2001] survey reported the cultural profile of each country in terms of a score and a ranking for each cultural dimension. This indicates that the strength of each side of each dimension can vary from user to user. For instance, users can be very high uncertainty avoidant, or marginally high uncertainty avoidant.

We have already shown that our test subjects scored in the low to medium range of the cultural dimensions tested (see Annexure E). We proposed therefore that performance levels may be significantly affected only if the users display high levels of a particular side of a cultural dimension. This aspect was not controlled for in the experiment.

The validity of this variable can help to overcome the need for trade-offs between conflicting dimensions as well as help to reduce the cost of accommodating subjective culture into user interface design. If performance is only affected when users display high levels of cultural dimensions, then it will not be necessary to accommodate for cultural dimensions that are low in

strength in the target user group. As a result, trade-offs between conflicting requirements of different dimensions are reduced, and the cost of accommodating for unnecessary dimensions is eliminated.

No explicit support for the validity of this variable has been found in the literature, and consequently further research needs to be conducted before it can be accepted as valid. Nevertheless, we have already presented theoretical evidence in Chapter 3 that subjective culture is related to usability from the perspectives of user acceptance, objective usability and context of use. As any changes to the user context of use can change the usability of the product [Bevan, 1995], it stands to reason that any variable relating to, or influencing the subjective culture of the user will also influence usability. As this variable influences the cultural profile of the user, we conclude that cultural dimension strengths should be considered for inclusion into the model.

5.2.2 Cultural Dimension Interplays

We suggested that one cultural dimension could override the impact of the other cultural dimensions on the interaction, particularly if the user displays a substantially high level of that one dimension. For example, a user that displays high levels of high uncertainty avoidance may find that high uncertainty avoidance dominates the interaction, thus reducing the effects of the other dimensions on the resultant performance. We did not control for this variable in the experiment. Like cultural dimension strengths, the validity of cultural dimension interplays can help to overcome design trade-offs and reduce costs.

No explicit support for the validity of this variable was found in the literature. However, Smith and Chang [2003] used Taguchi orthogonal arrays to study the impact of subjective culture on user acceptance. As the Taguchi method is used to reduce the possibility of dependent factors influencing each other [Dunckley and Smith, 2000], this implies that such interplays exist between dimensions.

The theoretical support in terms of the use of the Taguchi method together with the role that subjective culture plays in the context of use therefore leads us to the conclusion that cultural dimension interplays should be considered for inclusion into the usability model. Further research is however needed to establish the validity of this variable, as well as to more fully understand the extent of the interplays between cultural dimensions.

5.2.3 Relative Impact of Cultural Dimensions on Usability

This variable considers the possibility that the different cultural dimensions may have a stronger or weaker impact on usability. This variable was brought to light by the Smith and Chang [2003] study that reported differences in the relative impact of Hofstede's cultural dimensions on user

acceptability. As noted in section 2.6.2.1, they found that power distance contributed 22% to the variance in user preference, masculinity/femininity and individualism/collectivism had lower contributions of 9% and 2% respectively, and uncertainty avoidance virtually no contribution at 0.01%. This suggests that only some of the subjective cultural dimensions have a significant impact on usability.

The validity of this variable is again linked to the cost of accommodating for subjective culture into user interfaces. For example, if accommodating for a particular dimension enhances only one aspect of usability, the cost of accommodation may not be justified. This is particularly pertinent in the case where the measure that is enhanced is not the most important usability measure for the target user group (see section 5.6.1 and 5.8.1). In addition, this variable can help to alleviate conflicting interface requirements between cultural dimensions. It may be that a conflicting dimension does not affect usability, and therefore does not need to be accommodated into the design of the interface.

No further evidence to support or negate the validity of the relative impact of cultural dimensions has been found, and consequently the validity can only be based on the Smith and Chang [2003] study. However, as we identified variables (which we will explain in section 5.9.1), other than those tested that could have influenced the results of the study, additional empirical evidence is required before we can accept this variable as valid. Nevertheless, intuitively it is still possible that this variable could be found to be valid. This leads us to the conclusion that the relative impact of cultural dimensions should be considered for inclusion into the usability model.

5.2.4 Other Subjective Cultural Dimensions

Hofstede's [2001] set of cultural dimensions is one of many that have been put forward in the literature. The cultural models and related dimensions proposed by Victor [1992], Hall [1959] and Trompenaars [1993] were discussed in section 2.5.3. None of the additional subjective cultural dimensions identified in these models were controlled for in the experiment. We proposed therefore that these cultural dimensions could have influenced the test users' performance. In light of the potential for subjective cultural dimensions having stronger or weaker influences on usability (as discussed in section 5.2.3 above), it is also possible that these other subjective cultural dimensions could have overridden the impact of cultural dimensions that we tested in the experiment.

Two of Hofstede's subjective cultural dimensions are duplicated in the other cultural models. These are:

- Power distance, which is referred to as authority conception in Victor's model.
- Individualism / collectivism, which also appears in Trompenaars' model.

In addition, not all of the cultural dimensions identified in these models relate to subjective culture. Victor's model incorporates the objective cultural dimensions of language, social organisation, and environment and technology. Hall's model includes information flow and action chains, which we also concluded to be related to objective culture.

Consequently, we focused our attention on finding support for the validity of the additional subjective cultural dimensions included in these models. Some of these dimensions also appear in more than one cultural model. These are:

- Context, which is included in both Victor's model and Hall's model.
- Time, included in Hall's and Trompenaars' model; also referred to as temporal conception in Victor's model.
- Non-verbal behaviour, in Victor's model.
- Speed of messages and space in Hall's model.
- Universalism vs. particularism, neutral or emotional, specific vs. diffuse, achievement vs. ascription and environment in Trompenaars' model.

The literature provides some support for the validity of these additional subjective cultural dimensions. Massey et al. [2001] found that context, in conjunction with uncertainty avoidance and individualism/collectivism, epitomise the effects of culture on global virtual team communication. Hall's model of culture focuses on the way in which people communicate [Hofstede, 1996], which suggests that all of the dimensions in his model affect the usability of the interface as well. No further support for the validity of the other subjective cultural dimensions was found in the literature.

Due to the theoretical nature of the evidence that we found, we cannot accept the variables relating to other subjective cultural dimensions until further research provides empirical evidence of their validity. Nonetheless, the theoretical support, in conjunction with the influence of subjective culture on the context of use leads us to the conclusion that subjective cultural dimensions, other than Hofstede's, should be considered for inclusion into the usability model.

5.3 Variables relating to the Interface

Four variables relating to the test interface that need to be validated, were identified in Chapter 4 (section 4.7). The results of the statistical tests performed on the experiment data indicated that one of the two test interfaces in each set was generally more usable. We identified the following variables that could have caused this general increase in usability: partial representation of the cultural dimensions, usability principles, heuristics and guidelines, the relative impact of components, and the nature of the cultural dimensions. Based on the

evidence found in the literature, we concluded that two of these variables are valid for inclusion into the conceptual model of usability, and that there is sufficient theoretical evidence to consider the other two variables for inclusion as well. We will now discuss these four variables and the evidence provided in the literature in more detail.

5.3.1 Partial Representation of Cultural Dimensions

This variable considers the possibility that not all components of a particular interface reflect the same cultural profile. The variable came to light as a result of the cultural profile assessment of the test interfaces used in the experiment: the different components on each page of each website displayed different sides of a particular dimension. For example, the Barnes and Noble website (Annexure A-5.2) was evaluated as short-term oriented for the metaphors, navigation and interaction components, long-term oriented for the conceptual model component and both long and short-term oriented for the appearance component. This could result in users responding to different sides of the same dimension. We proposed therefore that partial representation could have distorted the experiment results. This aspect was not controlled for in the experiment.

Only one of the studies identified in the literature provided additional evidence that partial representation is a valid variable. In the Forer and Ford [2003] study, the cultural profile of each user interface component was assessed in order to identify appropriate test interfaces. During this identification process, numerous interfaces were evaluated and discarded, as it was found that the interface components were not consistently designed to display characteristics appropriate to specific sides of the dimensions. The websites that were finally chosen as test interfaces still contained components that did not display the required cultural dimensions.

Given the number of websites evaluated and discarded for the experiment and the Forer and Ford study, it is evident that the majority of existing web interfaces have not been specifically designed to accommodate subjective culture. We can therefore accept partial representation as a valid variable that needs to be included into the conceptual model of usability, if existing websites are chosen as test interfaces.

5.3.2 Usability Principles, Guidelines and Heuristics

This category of variables considers the effect of usability characteristics on the general usability of the interface. As discussed in section 2.4.5, numerous usability principles, guidelines and heuristics have been put forward in an attempt to enhance usability and increase user performance. We did not evaluate the test interfaces used in the experiment in terms of these usability characteristics. Consequently we proposed that the test interfaces that were identified as 'better' could have been generally more usable if the designers accommodated these usability characteristics into the design of the interfaces.

The literature that we investigated provides conflicting evidence for the validity of these variables. On the one hand, the developers of these principles, heuristics and guidelines advocate that they are applicable to most interfaces [Shneiderman 1998; Nielsen, 1993; Dix et al., 1998] and should therefore be included in user interface design. However, as discussed in section 2.4.2, interfaces are used within a specific context, and what is usable in one context is not necessarily usable in other [Preece et al., 2002]. Therefore, not all of these characteristics are relevant to every software product, so it is necessary to select those that are applicable to the user and the system [Newman and Lamming, 1995]. In addition, a number of researchers and practitioners suggest that such usability characteristics are inherently flawed because they do not take into account the context within which the product is being used, and because they are culturally situated. We discuss these flaws in more detail below.

5.3.2.1 Lack of Context

Principles such as those proposed by Dix et al. [1998], Shneiderman [1998] and Mayhew [1992] are based on theories of cognitive psychology [Mayhew, 1992], as discussed in detail in section 2.4.5. However, the cognitive approach focuses more on the syntactic and lexical features of the interface than the semantics [Vicente, in Hall et al., 2003]. Consequently, principles do not take into account a specific context, instead, they guide developers in the 'what' of design, rather than the 'how' of design [Newman and Lamming, 1995]. For example, Skinner's Behaviourism theory shows that people will perform better if they are rewarded [Fulton, 2002]. The principle of 'reward users often' can be derived from this cognitive psychology theory. Designers need to find the balance between too often and not often enough to provide a sufficient level of motivation for the user to continue using the system. The obvious question that arises is how often is often enough, and how often is too often? Skinner's theory may help designers to understand the broader concerns of users, but it will not tell the designer exactly how often to reward the user, or how to reward the user, within a specific context of use [Fulton, 2002]. As culture is part of the user context of use, principles should also provide design advice that is specific to the cultural context of the product's use.

Lower level characteristics such as guidelines and heuristics, which are based on principles, would therefore have the same inherent flaw [Hall et al., 2003]. We could not find any additional support for this in the literature, so we reviewed a set of heuristics to test Hall et al.'s supposition. We chose to use Nielsen's [1993] ten usability heuristics, as they are extensively used and frequently cited in international publications (for example, Dix et al., 1998; Preece et al., 2002), and have subsequently been translated into guidelines for web interface design and evaluation [Instone, 1999]. Our analysis provided further support that heuristics do not provide context-specific design advice. For example, the heuristic 'visibility of system status' advocates that users should be kept informed by providing appropriate feedback, within reasonable time'. 'Appropriate feedback' can take many forms, and is dependent on the context of use. Similarly,

'reasonable time' may be very different for the context of safety-critical and business critical systems [Kotonya and Sommerville, 1998]. From a cultural context perspective, we found two examples from Nielsen's [1993] heuristics where the acceptability of the measures is dependent on subjective cultural dimensions. These are:

- 'Aesthetics and Minimalist Design' proposes that the use of different colours should be minimized. The question that arises, from a detailed design perspective, is how many colours is enough, or too much? Relating this to a subjective cultural perspective, the answer to this question is dependent on the cultural profile of the user. For example, Marcus [2001] points out that for masculine users, the use of more than 3 colours would be too much, but for feminine users, this number would be too little. Hence, what is aesthetically pleasing to one side of a cultural dimension may not be pleasing for the other side of that same dimension, or for other dimensions.
- 'Visibility of system status' requires that appropriate feedback be given within reasonable time. For a high uncertainty avoidant user, 'appropriate feedback' should take the form of very precise, detailed and complete information that uses simple, clear and consistent terminology and images [Marcus, 2002]. In contrast, a shorter, more general message would be quite acceptable to low uncertainty avoidant users. Similarly, masculine and short-term oriented users want to complete tasks in the shortest possible amount of time. Consequently the amount of time that is considered reasonable by these users would be substantially shorter than for feminine and long-term oriented users.

5.3.2.2 Culturally situated

Although principles, heuristics and guidelines do not provide context-specific advice, they are inevitably developed by some person who is immersed in a particular culture [Hall et al., 2003]. This suggests that the cultural profile of the developers will be inherent in the usability characteristics that they develop, as well as in the user interface generated as a result of accommodating those characteristics into the design of the interface [Honold, 2000]. Consequently, the problem with usability characteristics is that they are culturally biased [Smith et al., 2004], and therefore do not accommodate context diversity.

The problem with culturally situated interfaces is that developers of usability characteristics, designers and users can belong to very different cultural universes [De Souza and Dejean, 1999; Mayhew, 1992]. As culture forms part of the context of use, this means that what may be usable to designers with a specific cultural profile, may not be usable to users with a different cultural profile [Bevan, 1995]. This problem is further exacerbated in that the extent to which an interface is evaluated to accommodate a heuristic is often determined by the cultural profile of the evaluator. Consider, for example, a web interface that incorporates breadcrumbs but not a site map. A low uncertainty avoidant evaluator would probably still rate the interface as high uncertainty avoidant. In contrast, a high uncertainty avoidant evaluator would probably rate the interface as low uncertainty avoidant.

Hall et al. [2003] presented theoretical evidence that the cultural profile of any set of principles is equivalent to the cultural profile of the person that developed those principles. They reviewed Shneiderman's [1998] eight golden rules in terms of the cultural model put forward by Trompenaars and Hampden-Turner [1997]. Hall et al. concluded that all of Shneiderman's rules are 'specific to the US culture that produced them' [2003, p 84].

Further evidence was found when we tested Nielsen's set of usability heuristics in a similar way. Using Hofstede's cultural model as a basis, we first identified Nielsen's probable cultural profile based on his country of origin and current country of residence. We then examined the heuristics to identify whether or not they displayed a particular cultural profile. Finally we compared the heuristics' cultural profile to Nielsen's cultural profile.

a. The Cultural Profile of the Developer

Nielsen was born in Denmark, obtained a PhD in user interface design/computer science from the Technical University of Denmark, and currently resides and conducts business in the USA. In terms of Hofstede's model of culture, the dominant cultural profiles recorded for these countries are as follows:

USA: Low power distant, individualist, masculine, low uncertainty avoidant and short-term oriented

Denmark: Low power distant, individualist, feminine, low uncertainty avoidant and short-term oriented

b. The Cultural Profile of the Heuristics

A detailed explanation of Nielsen's definition of his 10 heuristics and our analysis of their cultural bias is presented below.

- (1) Visibility of system status – the system should always keep users informed about what is going on, through appropriate feedback within reasonable time. 'Always keep users informed about what is going on' is a high uncertainty avoidance trait, in that it avoids uncertainty. As already explained above; 'within reasonable time' is vague and depends on the time-orientation and masculinity levels of the user.
- (2) Match between the system and the real world – the system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order. The 'real world' is dependent on the user's context, and is therefore specific to the user's objective and subjective culture. If this is achieved it is to cater for high uncertainty avoidant users, as it matches what is familiar to them.

- (3) User control and freedom – users often choose system functions by mistake and will need a clearly marked ‘emergency exit’ to leave the unwanted state without having to go through an extended dialogue. Therefore, ‘undo’ and ‘redo’ functions should be supported. This heuristic corresponds to masculine users as they strive for mastery of skills and require navigation oriented to exploration and control. Short-term orientation is also supported as this heuristic prevents users from having to go through extended dialogues.
- (4) Consistency and standards – users should not have to wonder whether different words, situations, or actions mean the same thing. Subsequently, platform conventions should be implemented to allow greater familiarity to users. This heuristic tries to reduce ambiguity by ensuring words or actions (platform convention) mean the same thing, thereby catering for high uncertainty avoidant users.
- (5) Error prevention – even better than good error messages is a careful design which prevents a problem from occurring in the first place. Preventing an error by design is to make a risk-free user interface that high uncertainty avoidant users require. Reducing possible errors creates the impression of user mastery, thereby also catering for masculine users.
- (6) Recognition rather than recall – make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate. ‘Visible and/or familiar objects with instructions for use’ will reduce uncertainty and is therefore related to high uncertainty avoidance. ‘Easily retrievable instructions’ caters to short-term orientation and masculinity as it reduces the time spent to complete the task. The utilitarian nature of objects and instructions will also increase user mastery, thus again catering to masculine users.
- (7) Flexibility and efficiency of use – accelerators, unseen by the novice user, may often speed up the interaction for the expert user such that the system can cater for both inexperienced and experienced users. Allow users to tailor frequent actions. ‘Tailored frequent actions’ is an individualist trait as it allows users to customize the system according to the way they want to work. The use of accelerators will appeal to short-term oriented users as they can increase their speed of performance. Flexibility and efficiency support exploration which is a masculine requirement. Novice and expert user distinction is also a status distinction related to masculine users.
- (8) Aesthetic and minimalist design – dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility. As discussed previously, ‘aesthetically pleasing’ is dependent on the users’ mental programming and is therefore specific to their cultural profile. From a more general perspective, however,

minimalist design requires ‘no irrelevant information to be included’ to avoid ambiguity (high uncertainty avoidance); extraneous information is to be avoided to prevent slow-down (short-term orientation and masculinity); and a minimalist design is utilitarian and therefore related to masculinity.

- (9) Help users recognise, diagnose and recover from errors – error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution. Precise, plain language error messages support high uncertainty avoidance. Error messages that constructively suggest solutions are low-power distant. Precisely indicating a problem accompanied by solutions is highly utilitarian and hence supports masculinity.
- (10) Help and documentation – even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user’s task, list concrete steps to be carried out, and not be too large. ‘Easy to search help’ supports high uncertainty avoidance users as it is unambiguous. ‘Task focused’ help supports masculinity and individualism. Providing ‘concrete steps to be carried out’ supports both masculinity and high uncertainty avoidance as it is certain and provides an executive view; providing ‘not too large’ help creates a short-term oriented design.

Table 5.1 summarises our assessment of the cultural profile of the heuristics.

Heuristic	Relevant Cultural Dimension
Visibility of system status	High uncertainty avoidance, masculine
Match between the system and the real world	Culturally specific
User control and freedom	Masculine, short-term oriented
Consistency and standards	High uncertainty avoidance
Error prevention	High uncertainty avoidance
Recognition over recall	Masculine, short-term oriented
Flexibility and efficiency of use	Masculine, individualist
Aesthetic and minimalist design	Masculine, short-term oriented
Help users detect, diagnose and recover from errors	High uncertainty avoidance, short-term oriented
Help and documentation	High uncertainty avoidance, masculine, short-term oriented

Table 5.1: Comparison of Nielsen’s Heuristics to Hofstede’s Cultural Dimensions

c. The Cultural Profile of the Developer and the Heuristics Compared

As Table 5.1 illustrates, the cultural profile of the heuristics, in order of most frequently occurring cultural dimension to least, is high uncertainty avoidant, masculine, short-term oriented, individualistic and low power distant. Nine out of the ten heuristics are appropriate for users who are high uncertainty avoidant, masculine and short-term oriented. High uncertainty

avoidance does not correlate to the dominant cultural profile of either of the two countries that have influenced Nielsen's cultural profile. Short-term orientation correlates to the dominant cultural profiles of both Denmark and USA, but masculinity correlates only to USA. It is interesting to note that according to Hofstede [2001] masculinity is dominant in many of the western countries, such as Great Britain, Canada, Germany and Australia.

It would appear that the theory that a set of usability characteristics will display the same cultural profile as its developer has been negated by the dominance of high uncertainty avoidance rather than low uncertainty avoidance in Nielsen's heuristics. However, high uncertainty avoidant sites are designed to reduce uncertainty, which naturally leads to increased performance in terms of accuracy, speed and satisfaction (see section 5.3.4 below). Common sense dictates therefore that high uncertainty avoidance is a better guideline for general usability than low uncertainty avoidance. The other two dominant dimensions of short-term orientation and masculinity are highly supported by the heuristics, which are consistent to Nielsen's expected cultural profile. Therefore, this theory is partly substantiated.

In summary, it has been contended that usability principles, heuristics and guidelines are inherently flawed because they do not take into account the context within which the product is used, they are not easily applicable to a specific case, and they are specific to one particular cultural profile [Hall et al., 2003]. This has in part been supported by other studies [Fulton, 2002; Honold, 2000; Smith et al., 2004], as well as by our own examination of Nielsen's usability heuristics. However, principles, heuristics and guidelines still serve a valuable purpose. They provide general guidance that is intended to inform interface design [Preece et al., 2002], provide designers with a more user-centred approach to the interface design [Fulton, 2002], and are valid for at least some cultural and other user contexts of use [Hall et al., 2003].

This leads us to conclude that usability principles, heuristics and guidelines are valid categories of variables that should be included into the conceptual model of usability. However, it will be necessary to identify and include only those that are relevant to the specific context of use for which the interface was intended.

5.3.3 Relative Impact of Components

This variable considers the possibility that the interface components of metaphors, conceptual models, navigation, interaction and appearance may influence usability to a higher or lesser degree. The type of interface could also affect the relative impact of the components. For example, ease of navigation on a website interface could affect performance significantly more than the appearance or the metaphors of the interface. In contrast, when using a traditional user interface (such as a word-processing or spreadsheet application), performance may be affected more by the conceptual model or appearance of the interface. The relative impact of

interface components was not controlled for in the experiment. We suggested therefore that the results of the experiment could have been affected by this variable, particularly if the test tasks did not cover these components equally between the interfaces.

Should this variable be found to be valid, it may not be necessary to accommodate each dimension into the design of each user interface component. This will once again reduce the costs associated with culturally sensitizing interfaces.

Some support for the validity of this variable was found in the Forer and Ford [2003] study. They reported different satisfaction ratings between the appearance, metaphors and interaction components of the test interfaces. However, the interface components displayed characteristics relevant to both sides of some of the cultural dimensions. Therefore these differences could have been attributable to differences in the interfaces' cultural profiles rather than a weaker or stronger influence of the components on satisfaction.

Consequently, further research is required before we can accept its validity. Nevertheless, the Forer and Ford study highlights the possibility that this variable could be valid. This leads us to conclude that the relative impact of interface components should be considered for inclusion into the usability model.

5.3.4 Nature of the Cultural Dimensions

This variable considers the possibility that the inherent characteristics of a specific side of different cultural dimensions provide a generally more usable interface than the opposing side of the same dimension. The interfaces used in the experiment evaluated to be high power distant, high uncertainty avoidant, masculine or collectivist were found to be the better sites, as depicted in Table 5.2.

Dimension	Side	Site
Power Distance	High	University of Malaysia (www.uum.edu.my)
Uncertainty Avoidance	High	Likouris Travel (www.travelagent.gr)
Masculinity vs. Femininity	Masculine	Sony (Sweden) (www.sony.se)
Individualism vs. Collectivism	Collectivist	Costa Rican National Park (www.tourism.costarica.com)

Table 5.2: Better Test Interfaces

We proposed in Chapter 4 that these interfaces were generally 'better' due to their incorporation of characteristics relevant to a particular side of a specific cultural dimension. For example:

- The increase in speed and accuracy levels obtained on the high uncertainty avoidant site could be attributable to the fact that high uncertainty avoidant sites are designed to reduce uncertainty. The design provides clear and familiar metaphors, simple, clear articulation and limited menu options, simple and limited navigation controls, precise and detailed feedback of status, simple and clear imagery and highly redundant coding [Marcus, 2002]. All these characteristics would naturally cater for more accurate and speedier completion of tasks. This could also have increased satisfaction levels as users would feel that the task had been accomplished quickly and correctly.
- Masculine site design incorporates similar characteristics to those of high uncertainty avoidant sites. For example, masculine sites are designed to provide limited navigation choices, and high-level executive views, and are goal and work-oriented [Marcus, 2002], thus providing for quick results of limited tasks. These characteristics would also naturally increase the speed and accuracy levels obtained, thereby possibly also increasing satisfaction levels.
- High power distant sites also provide limited navigation choices, and wizards or guides to assist with navigation [Marcus, 2002], thereby increasing the speed, accuracy and satisfaction levels obtained. However, it is noted that a significant result was obtained only in speed levels when compared to the low power distant site

In contrast to the high uncertainty avoidant, masculine and high power distant sites, the increase in satisfaction levels obtained on the collectivist site cannot be explained in terms of the characteristics of this side of the dimension.

The validity of this variable has important implications. Our original hypotheses were based on the belief that the cultural profile of the interface should be matched to the cultural profile of the users in order to enhance usability and performance, as proposed by Smith and Chang [2003]. However, if a particular cultural profile is found to increase the usability of interfaces for all users, this would invalidate this belief. Conversely, this would still provide evidence that Hofstede's cultural dimensions are related to usability, just not in the way that we originally hypothesized.

The validity of the nature of the cultural dimension is somewhat supported by one other study identified in the literature. As discussed in section 6.2.1, Smith and Chang [2003] reported that Chinese users preferred interfaces that displayed high power distant, high uncertainty avoidant, masculine and individualist characteristics. Other than the individualism/collectivism dimension, the preferred dimensions correlate to the findings of the experiment. At a superficial level, it could be stated that Smith and Chang's findings correlate to the findings of the experiment, and therefore the theory that interfaces that display the characteristics of certain sides of each dimension will be better interfaces, is supported. In addition, Smith and Chang expressed

surprise at the Chinese users' preference for sites that displayed individualism, in contrast to traditional perceptions of the Chinese as being a collectivist society. However, as we will explain in section 5.9.1, a number of possible variables were identified, other than the cultural dimensions tested, which could have directly influenced the results of their study. We therefore cannot accept the evidence provided by the Smith and Chang study to support the validity of this variable.

We also cannot accept the evidence provided by our own experiment, as all of the variables discussed in this chapter could have contributed to a larger or lesser extent to the results of the experiment. However, as two independent studies have brought to light this variable, we conclude that the nature of the cultural dimensions should be considered as a topic for further research.

5.4 Variables relating to User Acceptance

This category of variables relates to the existing body of knowledge surrounding the research area of user acceptance, and in particular to the Technology Acceptance Model. As discussed in section 2.6.2.1, the Technology Acceptance Model was one of many models developed in an attempt to explain why people accept or reject information technology [Davis, 1989]. It was established in Chapter 2 that user acceptance is a measure of user satisfaction, and is therefore a measure of subjective usability. User acceptance variables were not specifically controlled for in the experiment, thus possibly distorting the satisfaction levels of the test subjects. Consequently we proposed in Chapter 4 that any variable influencing user acceptance is a valid variable that influences usability.

The literature provides conflicting evidence of the validity of these user acceptance variables. On the one hand, numerous studies have verified the validity of the model as a way of explaining and predicting user acceptance. On the other hand, some studies have highlighted that the model is not applicable to all contexts of use. We discuss this evidence in more detail next.

Since its inception, several studies (as listed in Table 5.3) have validated the Technology Acceptance Model's ability to explain why users accept or reject information technology. These studies have also identified many additional variables that influence user acceptance. Some of these variables directly influence user attitudes, whilst others indirectly influence attitude through the constructs of ease of use and usefulness. A selection of these variables is presented in Table 5.3.

Variable	Explanation of Variable	Study
IT Role	Role with regards to information technology, for example novice or expert user.	Agarwal and Prasad [1999]
Education level	Level of general education, educational qualification, eg high school, postgraduate degree.	Agarwal and Prasad [1999]
Prior similar experience	Experience with similar applications.	Agarwal and Prasad [1999]
General computer self-efficacy	The user's confidence in their ability to use technology.	Agarwal et al. [2000], Brown [2002], Venkatesh [2000], Venkatesh and Davis [1996]
System self-efficacy	The user's confidence in their ability to use the particular language or mode of interaction of a given system.	Agarwal et al. [2000]
Organisational support	Encouragement by top management and allocation of adequate resources.	Anandarajan et al. (2002)
Computer Anxiety	The ability to succeed with a new system.	Brown [2002], Venkatesh [2000], Venkatesh and Davis [1996]
Ease of Finding	Ease of navigation and allow easy return to previously displayed pages.	Brown [2002], Lederer et al. [2000]
Ease of Understanding	Understandable and consistent graphics and terms.	Brown [2002], Lederer et al. [2000]
Tool functionality	Functionality provided by the tool.	Dishaw and Strong [1999]
Tool experience	Individual abilities to use tools	Dishaw and Strong [1999]
Task-technology fit	Match between user task needs and available functionality of the IT.	Dishaw and Strong [1999]
Information Quality	The accuracy, timeliness, relevance and completeness of information.	Lederer et al. [2000]
Facilitating Conditions	Availability of support staff to help users to overcome barriers to technology usage, during the early stages of learning and use.	Venkatesh [2000], Venkatesh and Davis [1996]
Computer Playfulness	The degree of cognitive spontaneity in microcomputer interactions	Venkatesh [2000]
Perceived Enjoyment	The extent to which the activity of using a computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated.	Venkatesh [2000]
Objective Usability	Time taken to complete a task as a ratio of time taken by experts to complete the task.	Venkatesh [2000]
Professional status	The extent to which using the system will enhance the status of the user as perceived by the user's colleagues.	Succi and Walter [1999]
Cultural design preferences	Different input and output devices and interaction styles.	Evers and Day [1997]

Table 5.3: Examples of Variables that Influence User Acceptance

Further support for the validity of these user acceptance variables was found when comparing them to the context of use variables. A number of duplications and correlations were found. We

will defer the discussion of the correlations to Chapter 6, where we present a revised and expanded context of use as a basis for the conceptual model of usability. A summary of the duplications is listed in Table 5.4.

User Acceptance Variables	Duplications in Context of Use
IT Role	User Role
Education Level	Qualifications
Prior Similar Experience	System knowledge / general computer experience
Organisational Support	System use
Tool Functionality	Major functions
Facilitating Conditions	Assistance available

Table 5.4: Duplicated and Correlated User Acceptance Variables

Therefore, even if our argument (presented in section 2.6.2.1) that contends that acceptance is a measure of usability is rejected, the user acceptance variables that have been shown to be duplications and/or dimensions of the context of use variables, can still be accepted as valid, as the context of use variables have been shown to be irrefutably linked to usability.

In contrast to the above evidence that supports the validity of the user acceptance variables, other studies [Evers and Day, 1997; Anandarajen et al., 2002; Brown, 2002] have shown that the Technology Acceptance Model, and the variables influencing the model's constructs, are not valid for all contexts of use, in particular, the user's subjective culture, technological environments, specific types of software applications, and the user's nationality and ethnicity.

The Evers and Day [1997] study identified differences in user acceptance flows (variables that influence usage) between nationalities and ethnic groups, suggesting that perceived ease of use and usefulness are not always the main determinants of user acceptance for every nationality or ethnic group. They reported that Chinese users will try to work with a perceived useful interface even if it is hard to use. In the Chinese acceptance flow, preferences also directly influence satisfaction, which implies that the users' demands for ease of use are met when the interface looks the way they want it to. Usefulness also influences satisfaction, therefore external design features are not enough to satisfy the Chinese – certain functionality levels must also be met. In contrast, Indonesians will tend to give up if an interface is hard to understand. This is an interesting, but understandable finding given that Hofstede [2001] reported that Indonesians generally have higher uncertainty avoidant levels than Chinese. The acceptance flow also implies that when Indonesians find a system easy to use, they expect to be happier in using the system. It seems that being genuinely happy with a system is less important to Indonesians than the knowledge that it will be easy to use. In contrast to both Chinese and Indonesian users, the acceptance flow process modeled for Australian users implies that if preferences are met then the users will be satisfied, irrespective of how easy to use or useful the system is. These acceptance flows are summarised in Table 5.5.

Nationality / Ethnicity	Acceptance Flow Process
Chinese	Preference → usefulness → satisfaction → behaviour
Indonesians	Preference → ease of use → satisfaction → behaviour
Australians	Preference → satisfaction → behaviour

Table 5.5: Acceptance Process Differences Between Chinese, Indonesian and Australian Users

The differences in acceptance flows reported by Evers and Day [1997] suggest that the contribution of perceived ease of use and usefulness to user acceptance differs between nationalities. As shown in Table 5.5 for example, perceived usefulness influences Chinese user acceptance, but not perceived ease of use. Consequently the variables that influence perceived ease of use do not need to be controlled for, as perceived ease of use does not influence acceptance for Chinese users. Thus the validity of these acceptance variables is dependent on the nationality of the users. User nationality and ethnicity are discussed in more detail in section 5.7.

Anandarajen et al.'s [2002] findings have similar implications, but in terms of subjective culture. They reported that, in a collectivist culture, for example, the variables influencing usefulness would not influence usage because usefulness itself does not influence acceptance in this context. In addition, perceived enjoyment had no direct effects on attitudes, and individually based training did not enhance computer skills and therefore had no effect on perceived ease of use. However, organizational support is positively related to ease of use in a collectivist society. These findings indicate that the validity of these variables is dependent on the subjective cultural profile of the users.

Brown [2002] reported that for users in developing countries, ease of finding and ease of understanding both influenced perceived ease of use; however, ease of understanding had a stronger influence than ease of finding. This shows that it is more important for the users to understand the interface, particularly if the language used in the interface differs to the home language of the users. In addition, user characteristics of computer self-efficacy and computer anxiety influenced perceived ease of use. This suggests that in the context of developing countries, where the social, educational and economic environments are not conducive to widely accessible technology, computer anxiety and self-efficacy could be very important and highly influential to the user acceptance process.

Furthermore, the variables identified to influence perceived ease of use and usefulness are only relevant within the context of specific types of applications. For example, computer self-efficacy was found to influence perceived ease of use for a variety of applications, but not for spreadsheets [Lederer et al., 2000].

The evidence discussed above has shown that the validity of the user acceptance variables is dependent on the subjective culture and nationality of the user, as well as the task context. Therefore we can conclude that when investigating the relationship between culture and usability it is only necessary to identify and control for the acceptance variables relevant to the user and task contexts specific to the context of the research. This provides opportunities for additional research to further define these variables and their validity within specific cultural and task contexts.

5.5 Variables relating to Speed of Performance

Four variables directly influencing speed of performance were identified from the literature. These are:

- (1) Hardware platforms, which relate to the technical environment of the software, and includes aspects such as processing speed and memory capacity.
- (2) The number of navigational decisions, which relates to the number of decisions that users have to make in terms of which links to follow to find the required information.
- (3) The number and length of bodies of text, which relates to the number and length of bodies of text that need to be read in order to find the information required for the task.
- (4) The level of Internet traffic, which relates to the number of users accessing the Internet at any particular time.

Only one study [Forer and Ford, 2003] suggested that it was necessary to control for these variables to ensure that variability of speed was not accidentally affected. Additional support for the validity of the hardware platforms and levels of Internet traffic was found by comparing them to the context of use variables (which were discussed in section 2.4.2). Hardware platforms can be seen as a dimension of the hardware variable listed in the context of use breakdown in both the equipment context as well as the technical environment category of the environment context. Similarly, the level of Internet traffic can be seen as an additional dimension of the technical environment category. As both of these variables can be related to variables that have been accepted as valid, their validity can be accepted from a theoretical perspective.

Test tasks are generally designed in a way that requires users to look for information on different pages of a particular website [Spool et al., 1999]. Accessing the information requires users to select and navigate through appropriate links, and to read through bodies of text. Neither of these variables was controlled for in the experiment. This could have resulted in the following scenario:

- For the first interface in each set, the tasks required 3 navigational decisions and reading through 2 bodies of text containing 10 lines each.

- For the second website in each set, the tasks required 5 navigational decisions and reading through 4 bodies of text containing 20 lines of text.

It is evident from the above example that variations in these variables between the interfaces could have substantially affected the speed with which the users could complete the tasks set for each interface.

Notwithstanding the limited support that the literature has provided for the validity of these four variables, it is difficult to refute their validity, as they all plainly influence speed of performance. Consequently, speed of performance variables are considered to be valid for inclusion into the usability model.

5.6 Variables relating to Objective Culture

Objective culture relates to those characteristics of the user that are visible and tangible, and includes different levels such as national, regional, gender and social class. This was discussed in section 2.5.2.4. Numerous studies investigating the relationship between objective culture and usability have been conducted [for example, Bourges-Waldegg and Scrivener, 1998], resulting in substantial empirical and theoretical evidence that objective culture influences the usability of computer-based systems. Although many of the variables relating to objective culture have already been discussed in Chapter 2, the literature investigation has brought to light that the nationality and ethnicity of users are important variables that should be included into the usability model.

Although the concepts of nationality and ethnicity were discussed previously in section 2.5.2.4, we provide a more detailed definition of these terms below to provide the reader with a better understanding of these concepts and how they influence usability.

- Nation: The term 'nation' is often used to mean 'nation-state', which is a sovereign state with its own government, boundaries, defence forces, etc., and symbolic markers of nationhood such as a flag, an anthem, local currency and a head of state [Hartley, 2002, p 150]. Most nations are multi-racial, multilingual and multicultural to some degree, for example, the United States of America and South Africa
- Nationality: the status of belonging to a particular nation by birth or naturalization [Wordnet, 2003].
- Ethnicity: 'of peoples from other cultures' [Hartley, 2002, p 83]. Members of an ethnic group differ with regard to certain cultural characteristics from other members of their society [Theodorson and Theodorson, 1969], in terms of racial, national, religious, linguistic, or cultural heritage [Wordnet, 2003]. A person can therefore be a member of a

particular ethnic group, especially belonging to a national group by heritage or culture but reside outside its national boundaries

In a situation where, for example, a Chinese person was born in, and permanently resides in, America, the above definitions would define such a person as being of Chinese ethnicity but with an American nationality.

The literature investigation has identified that nationality and ethnicity influence usability indirectly by influencing:

- (1) the validity of user acceptance variables,
- (2) the relative importance of usability measures,
- (3) user preferences, and
- (4) cognitive abilities of the users.

The effect of nationality and ethnicity on user acceptance variables has already been discussed in section 5.4. Consequently, we discuss the latter three variables below.

5.6.1 Relative Importance of Usability Measures

This variable considers the possibility that one or more of the usability measures are more important than the others. Given that subjective culture forms part of the user context of use, we proposed in Chapter 4 that subjective culture could also influence the relative importance that users place on the usability measures. For example:

- high uncertainty avoidant users may perceive accuracy to be more important than speed or satisfaction;
- individualist users may rank satisfaction as more important, given their focus on personal goals, and
- short-term oriented and masculine users may value speed of performance more due to their inherent need for quick results.

The influence of subjective culture on the relative importance of usability measures was tested in a study conducted by Vohringer-Kuhnt [2003]. He reported no significant correlations between these variables, suggesting that Hofstede's subjective cultural dimensions do not influence the relative importance placed on usability measures by users. However, his findings suggested that objective culture, and specifically nationality, results in significant differences in user perceptions of the relative importance of usability measures.

Test subjects in the Vohringer-Kuhnt study were sourced from different nationalities. Subjects were asked to rate the importance of a number of definitions of usability. These definitions were then correlated to one or more of the three usability measures of effectiveness, efficiency and satisfaction. Table 5.6 depicts the definitions and correlations that were discussed in Vohringer-Kuhnt's research report.

The relative importance of the measures for each nationality was computed based on the number of definitions relevant to each usability measure, and how important they were to the users. The findings are summarized in Table 5.7.

Definitions	Effectiveness	Efficiency	Satisfaction
Minimal resources		X	
Minimal effort		X	
Reaching Goals	X	X	
Learnability	X	X	
Error resistance	X		X
Feedback			X
Navigation	X	X	X
Completing tasks	X		
Wording			X
Design			X
Hedonistic quality			X
Accuracy	X		

Table 5.6: Definitions of Usability Measures

Nationality	Most Important	Important	Least important
American	Minimal effort, using minimal resources	Reaching goals, learnability, navigation	Feedback, error resistance
German	Navigation	Reaching goals and completing tasks, wording and design	Learnability, hedonistic quality
Australian	Navigation	Minimal resources and effort Completing tasks Hedonistic quality	
British	Reaching goals, navigation	Cognitive environment	Minimal resources, hedonistic quality
South African	Navigation	Minimal resources, design	Completing tasks, accuracy
Dutch	Reaching goals	Navigation Hedonistic quality	Overall quality and design

Table 5.7: Relative Importance of Usability Measures By Nationality

Table 5.7 shows that South African users rated efficiency and satisfaction to be more important than effectiveness; Americans rated efficiency to be more important, followed by effectiveness, whilst satisfaction is the least important measure. In contrast, Dutch users rated effectiveness and satisfaction as more important than efficiency. The reported differences in these importance ratings indicate that the relative importance of the measures of usability differs between nationalities, rather than between subjective culture. Thus, nationality influences the relative importance of usability measures.

5.6.2 User Preferences

Nationality has been reported to influence user preferences in terms of specific interface design features. Evers and Day [1997] conducted a study to examine users' culturally specific interface design preferences and the consequences of satisfying or not satisfying these preferences on user acceptance. They reported significant differences in specific interface design features between Asians and Australians, as summarized in Table 5.8.

User Interface feature	Percentage that liked the feature	
	Asians	Australians
Lots of different colours	73	29
Soft colours	68	32
Pull down menus	54	63
Fixed menus	54	31
Text based interfaces	43	8
GUI interfaces	64	78
Mouse as input device	89	72
Joystick	66	21
Touch Screens	73	47
Sounds	87	65
Instruction of machine through detailed commands	53	34

Table 5.8: Asian and Australian User Design Preferences

The preferences summarized in Table 5.8 suggest that in contrast to Australians, Asians prefer lots of different colours, soft colours, and the ability to instruct the computer through detailed commands. Asians and Australians both like the mouse as an input device best, with the touch screen second; however, more Asians like the mouse and the touch screen than the Australians. The data (in addition to those reflected in Table 5.8) also suggest that there are preference differences within Asian groups, specifically between the ethnic groups of Indonesians and Chinese. Indonesians like soft colours, black and white displays and pop-up menus more than the Chinese. Indonesians seem to like new technology and alternative input and output more than the Chinese. In contrast, Chinese like the use of many different colours more than the Indonesians. These findings indicate that there is a dominant set of user interface design feature preferences for each nationality and ethnic group. This provides further

evidence that nationality and ethnicity influence usability, by influencing the subjective usability measure of satisfaction.

5.6.3 Cognitive Abilities

Ethnicity has been reported to influence the cognitive abilities of users [Choong and Salvendy, 1998], as summarized in Table 5.9. Choong and Salvendy [1998] found that Chinese users performed better when using pictorial icon displays in contrast to Americans, whose performance increased when using textual displays. As illustrated in Table 5.9, the study reported that the verbal fluency of Americans is higher than the Chinese, whilst the visual ability of the Americans is lower than their Chinese counterparts. Visual ability can be seen as a dimension of the user's physical capabilities and limitations, which form part of the user context of use, and therefore influences usability. Verbal fluency can be seen as a dimension of linguistic ability, also identified as a variable relating to the user context of use. Therefore, the increase in performance reported for the American users using textual based interfaces is evidently linked to the cognitive capabilities inherent in this nationality. The same holds true for the increase in performance of Chinese users when using pictorial-based interfaces.

Cognitive Variables	Americans	Chinese
Cognitive style	Inferential-categorical	Relational-contextual
Verbal fluency	Higher	Lower
Visual ability	Lower	Higher
Digit span in STM	4.6 – 7.2	5.9 – 9.2

Table 5.9: Variables Relating to Nationality that Influence User Performance

In addition, differences in digit span in short-term memory and cognitive style were also reported in the Choong and Salvendy study. Cognitive style has already been shown to influence usability as it is included in the determinants of user performance as a dimension of the psychological characteristics of the user. Digit span in short-term memory relates to memory, which, as discussed in section 2.3.2 is a cognitive resource in the human information processing system, and is therefore also a determinant of user performance. Furthermore, we noted in section 2.4.2.1 that cognitive abilities also form part of the user context of use, which in turn influences usability. This suggests that ethnicity indirectly influences usability through its influence on the cognitive abilities of the users.

In summary, nationality and ethnicity influence the importance that users place on usability measures, the validity of user acceptance variables, the interface features preferred by users, and the cognitive abilities of the users. Nationality and ethnicity are therefore accepted as valid variables that need to be included into the usability model.

5.7 Performance Determinants

Several variables that are external and internal to the user were identified as determinants of user performance and were discussed in detail in section 3.3.4. These performance determinants have been included in the two contexts of use proposed by Bevan [1995] and Kirakowski and Cierlik [1999], providing additional support for their validity. Consequently we accept these variables as valid for inclusion into the model of usability.

Although we controlled for the more obvious determinants such as age, home language and computer skills in the experiment, we did not think it was necessary to control for all the performance determinants. In hindsight, and in light of our better understanding of the nature of these determinants, it is very likely that the performance of the users during the experiment was distorted as a result of these omissions. In this section we discuss how the omitted performance determinants could have influenced the results of the experiment.

5.7.1 Psychological Characteristics

Differences in cognitive styles were not controlled for in terms of the test subjects or the test interfaces used in the experiment. It is possible therefore that test subjects had varying cognitive styles that affected their performance. It is equally possible that the test interfaces accommodated different cognitive styles, thus adding to the impact of cognitive styles on the users' performance levels.

5.7.2 Knowledge and Experience

As discussed in section 4.4.2.1, test subjects were filtered in terms of the qualification that they were registered for, typing skills, general level of computer experience, length of time using a computer and prior experience with some of the test interfaces. However, it is possible that due to their prior experience, the test subjects' preconceptions of the test interfaces affected their performance. For example, test subjects that had prior experience with websites of the same genre could have expected the test interfaces to have similar functionality and layout to the ones that they had experienced previously. These preconceptions may also have arisen as a result of the users' real world experiences. For example, users who had purchased an airline ticket using a travel agent may have been expecting an interface with a similar conceptual model.

It is equally possible that a lack of real world experience could have affected performance. Test subjects could have experienced difficulty in anchoring signals because they did not have the domain knowledge required for using the test interfaces. For example, test subjects that had never purchased an airplane ticket would not be familiar with the process or the jargon required.

The filtering process used to identify test subjects did not cater for prior experience on website genres or domain knowledge for all the tasks. Therefore, the users' prior experience could have affected the cognitive processes of identification, analysis and response, resulting in lower performance levels.

5.7.3 Job and Tasks

Although the test subjects' experience was controlled for in general terms, their experience with specific interaction styles and input devices was not. In addition, the tools themselves afford different response times, particularly in terms of the size of the target and the distance that the target has to be moved. The use of the different interaction styles and input devices by the test interfaces was not controlled for. Therefore it is possible that some of the interfaces used interaction styles and input devices that were conducive to increased speed, or were consistent to the users' prior knowledge and experience.

5.7.4 Physical Characteristics

We controlled for colour-blindness in our experiment, but not for handedness. Left and right-handedness can influence the speed of performance due to the layout of the keyboard. In addition, ambidexterity can increase speed of performance, as left-handed users will not need to alternate between using the keyboard and a pen to write their answers down. Consequently, the speed of performance achieved during the experiment could have been affected by differences in handedness and ambidexterity of the users.

5.7.5 Physical Environment

All test subjects performed the test tasks in the same computer laboratory. Therefore, no further variables relating to the physical environment that could have distorted the results of the experiment are identified.

5.8 Impact of Variables on Experimental Design

The variables identified and discussed in this chapter have some implications on the way in which the experiment was conducted. These implications include the relative importance of usability measures, compulsory participation and time limits, adaptation of the cultural profile questionnaire, and the order in which the test subjects were exposed to the experimental test tasks. Each of these are discussed in more detail below.

5.8.1 Relative Importance of Usability Measures

In the experiment, we measured usability in terms of accuracy, speed and satisfaction. If a significant difference was found for any one of these measures, the hypothesis that subjective culture influences usability would have been accepted. However, if the relative importance of one usability measure is higher than the other two measures, this would mean that the overall usability cannot be determined simply by aggregating the usability levels for each measure. Instead, it will be necessary to weight the levels according to the relative importance attached to each measure by the users. This aspect was not controlled for in the experiment.

Although our initial proposal that subjective culture influences the relative importance of usability has not been supported, we have shown that the variable is still valid, but from an objective culture and task perspective. Furthermore, the task context influences the relative importance of these measures, as discussed in section 2.4.2.2. Preece et al. [2002] point out for example, that in the games context, the more effort required to complete a specific task will conversely result in a more enjoyable and fun experience. In contrast, when dealing with work-oriented software and interfaces, the accuracy and speed with which the task is completed could be more important than having fun. This provides further support for the validity of this variable, as it once again suggests that it will be necessary to determine the relative importance of the measures within the context of the task, and then to weight the usability measures accordingly.

This leads us to conclude that the relative importance of usability measures is a valid variable that should be included into the usability model. Further research is required, however, to establish the relative importance of the measures for each nationality and each type of task.

5.8.2 Compulsory Participation and Time Limits

Participation in the experiment was compulsory for all students enrolled for the 3rd-level course in Business Information Systems. Furthermore, time limits were imposed on each test task. As test subjects' responses are influenced by transient personal factors such as mood and motivation [Ghauri and Grønhaug, 2002], we suggested that the compulsory nature of the experiment, together with the time limits imposed, could have influenced the mood and motivation levels of the participants, thereby influencing their performance.

We have shown that motivation has been identified as a determinant of user performance and can be influenced by fear or interest (section 2.3.4.1). Too little or too much interest negatively affects motivation [Mayhew, 1992]. Compulsory participation, with little interest on the user's part, could have led to low levels of motivation. If however, the test subjects did consider the tasks to be important, then the time limits that were set for each of the tasks could have resulted in users increasing their speed in order to complete the tasks within the specified time limits.

In this case, accuracy could have been reduced due to the increase in user speed. On the other hand, the compulsory nature of participation did not stipulate minimum performance levels. This could have led the users to consider the task to be of low importance, also resulting in reduced performance levels.

It is evident that compulsory participation and time limits influence motivation levels in one way or another. As motivation is a known performance determinant, this leads us to the conclusion that compulsory participation and time limits should be avoided when conducting usability studies.

5.8.3 Adaptation of Cultural Profile Questionnaire

The test subjects used in the experiment were students who had little or no working experience. As a result, we adapted Hofstede's cultural questionnaire (Value Survey Model) to suit the context of the test subjects. Care was taken to identify appropriate questions through the use of the literature as well as a pilot study. However, the inconclusive results of the experiment led us to the possibility that the test subjects' cultural profiles could have been inaccurately evaluated. In particular, it is possible that users identified as low power distant could have been high power distant. Users who are high power distant would have answered positively phrased questions positively, and vice versa, as they may not have wanted to be seen to disagree with the question (and the researcher was at the time of the experiment also their lecturer). The majority of the power distance questions were phrased positively, where a positive answer indicated low power distance.

Due to the inherent characteristics of high power distant users we conclude that the adaptation of the cultural profile questionnaire could have influenced the results of the experiment. Consequently, we suggest that should existing questionnaires need to be adapted, it is important to take into consideration the cultural profile of the users, particularly with respect to high power distant users.

5.8.4 The Order Effect

The order effect considers the effect (on user performance) of the order in which the test subjects were exposed to each of the test interfaces. As noted in section 2.3.2, memory is affected by the serial position effect, which suggests that the first and last items in a list are remembered better than the items in the middle. Two opposing laws relevant to the serial position effect have been identified, namely the law of primacy [Lund, 1925 in ciadvertising.org] and the law of recency [Cromwell, 1956, in ciadvertising.org]. The law of recency suggests that the last material, still fresh in the person's mind through its recency, will be best remembered. In contrast, the law of primacy suggests that the earlier part of a list of items tends to be better remembered than the latter part [CogSci Dictionary, 2004].

In the context of experimental research, the laws of primacy and recency form the basis for the order effect and the carryover effect. The carryover effect is defined as 'the effect of previous trials on a subject's performance on subsequent trials', whilst the order effect is defined as 'the effect on performance attributable to the order in which treatments are administered' [Howell, 1989]. Howell notes that the primary disadvantage of related measures experimental design is that there may be either an order effect or a carryover effect from one session to the next. Essentially, these effects can be seen as increasing the user's knowledge and experience, as the experience gained from participation in the first test will be used by the user in the second test.

As discussed in section 4.4.4, we used a related measures (within-subjects) design for our experiment. Of the four sets of interfaces tested for differences, the second site in three of the four sets was found to have a higher level of general usability than the first site, suggesting a carryover or order effect on user performance. In terms of the carryover effect, the experiment was conducted over five sessions, therefore test subjects' performance in the earlier sessions could have affected their performance in later sessions. In addition, two test interfaces were used in each session. Therefore, an order effect could have occurred between the performance achieved on the first and second interfaces used in each session.

Some additional evidence for the validity of the order effect was found in the literature. McFarlane [1998] points out that there is a large primacy effect that affects the subject's ability to learn appropriate strategies for completing tasks in different treatment conditions. He reported that his test subjects formed 'rigid task strategies based on whatever treatment they saw first' [p 207]. This indicates a primacy effect because of how they were first told how to manage interruptions. For the experiment (Chapter 4) this could mean that after learning how or where to find the information from their interaction with the first website, the test users transferred this knowledge to the second website, thus finding it easier to locate the required information. On the other hand, if the second test interface used a different navigational design to the first, then the primacy effect would have reduced performance levels because the test subjects would have assumed, incorrectly, that the same strategy could be used in the second site. Either way, the law of primacy would have affected speed and accuracy, as well as satisfaction levels, particularly if satisfaction is a function of speed and accuracy. In addition, Smith [2004] confirmed the validity of this variable by stating that 'the users would have met the sites in random order to avoid bias' in the study performed by Smith and Chang [2003]. Therefore when conducting an experiment of this nature the order effect should be controlled for by randomizing the order in which the test subjects are exposed to each test interface.

5.9 Impact of Variables on Prior Studies

The variables discussed in this chapter impacts on the validity of the results reported in some of the literature sources used in this research. We review the methodologies used in two of the studies that provided extensive evidence for our research. As we will show, many of these variables were not controlled for in these studies. Some of the omitted variables could have influenced the results of these studies, thus raising doubts as to the validity of the evidence presented by the two studies. Consequently, cultural influences on usability and performance remain confused, necessitating further research. This provides further support for the need to establish a more detailed and robust conceptual model of usability that will help to increase the reliability and validity of the results from research of this nature.

5.9.1 The Smith and Chang Study

As discussed in section 2.6.2.1, this study focused on identifying the relative importance of Hofstede's [2001] dimensions on web site acceptability. Smith and Chang [2003] performed two sub-studies:

- Sub-study A investigated the influence of power distance, individualism/collectivism and uncertainty avoidance.
- Sub-study B investigated the influence of power distance, individualism/collectivism and masculinity/femininity.

Using Taguchi orthogonal arrays, four web interfaces were identified for each study. Each interface displayed appropriate characteristics of each side of each dimension, as identified by the Taguchi arrays. For each study, Chinese test subjects were asked to perform a series of tasks on the websites. After accessing the sites, the users were asked to complete a satisfaction questionnaire which incorporated questions relating to appropriateness of layout and navigation, ease of information access, level of trust engendered and likeliness of return and recommendation to others [Smith et al., 2004]. The results of the questionnaires were analysed to identify the preferred side for each dimension, and the percentage contribution made by each tested dimension. The results from the sub-studies were then combined to determine the relative importance of the four dimensions to user acceptance.

Two of the variables discussed in this chapter that were omitted and could have influenced the results of the study are:

- (1) Cultural dimension interplays – Although the test interfaces were evaluated in terms of three of Hofstede's dimensions, it is not clear whether the other two dimensions of each site were evaluated and controlled for. For instance, the four test interfaces that were used in sub-study A were evaluated to display the cultural dimensions as reflected in Table 5.10.

SUB STUDY A			
Interface 1	High power distant	Individualist	High uncertainty avoidant
Interface 2	High power distant	Collectivist	Low uncertainty avoidant
Interface 3	Low power distant	Individualist	Low uncertainty avoidant
Interface 4	Low power distant	Collectivist	High uncertainty avoidant

Table 5.10: Cultural Dimensions Tested in Sub-Study A [Smith and Chang, 2003]

Table 5.10 shows that the masculinity / femininity and time orientation characteristics of each of these interfaces were not evaluated or controlled for. This suggests the possibility that the users' preferences for high power distant, individualist and high uncertainty sites could have been attributable to the presence of masculine / feminine or short-term / long term characteristics in the test interfaces.

- (2) Relative impact of interface components – It is not clear whether the tasks set for each website in each study incorporated the same interface components. In addition, it is not clear whether the evaluation of the test interfaces assessed each interface component in terms of the cultural dimension characteristics that they displayed. It is therefore possible that user preferences were influenced by the relative impact of different interface components, which could have been further exacerbated by differences in the dominant cultural profiles of these components.

The Smith and Chang study also did not identify the cultural profile of the test users [Smith email]. Instead, users were chosen based on their nationality (Chinese). Smith and Chang [2003] expressed surprise that their test subjects preferred a web site displaying individualism characteristics, as this preference is in contrast to that expected from a collectivist culture such as the Chinese. However, as the cultural profiles of the test users were not identified, it is possible that the dominant cultural profile of the test subjects was different to the country's cultural profile, which could explain the surprising results for the individualism/collectivism dimension. In addition, the significance percentages of power distance between sub-study A and sub-study B were markedly different, as were the significance percentages for the masculinity/femininity dimension. This differentiation could be as a result of differences in the cultural dimension strengths of the users. For example, the user group in study A could have exhibited very high levels of high power distance in contrast to the user group in study B.

Although this does not influence the results of the study in terms of the relative impact of the cultural dimensions on user acceptance, it does not provide reliable and robust evidence that matching the cultural profile of the interface to the cultural profile of the users increases acceptance and usability. As the cultural profiles of the test subjects were not identified, there is little evidence that matching the cultural profile of the interface to the users increases

acceptance.

Other literature sources that we have used in this research (specifically in section 2.6.2.1) also did not identify the cultural profiles of the test users [Straub et al., 1997; Anandarajen et al., 2002; Massey et al., 2001]. This omission is surprising, as although Hofstede [2001] identified a dominant cultural profile for each country, not every citizen of that country has the same cultural profile. Cultural boundaries do not necessarily coincide with national boundaries [Duncker, 2002]. National states often comprise multiple cultures and ethnicities, such as those comprising South Africa. Therefore, it is very likely that a selected sample of citizens of Japan, Guatemala and South Africa would have the same cultural profile.

One of the criticisms against using cultural dimensions is that such models tend to stereotype users, as discussed in section 2.6. However, perhaps it is the researchers who are using such models that are doing the stereotyping, rather than the developers of the models themselves: omitting to identify the cultural profiles of the individual test users is a prime example of such stereotyping.

Failing to identify the cultural profile of the test users therefore raises doubts as to the validity of the results reported by these studies. For example, the Anandarajen et al. [2002] study reported that the perceived usefulness of information technology would not affect user acceptance for users who are high uncertainty avoidant, high power distant, collectivist and masculine. We interpreted these findings as evidence that there is a relationship between subjective culture and usability, and furthermore, that the influence of some user acceptance variables is dependent on the subjective cultural profile of the users. However, given that the cultural profiles of the test subjects were not identified, there is actually little evidence that the reported findings are valid for the cultural dimensions indicated.

The Straub et al. [1997] study suffers from the same flaw. The study provided evidence that users who are high uncertainty avoidant, high power distant, collectivist and masculine would reject communication media that is not information rich or does not support social presence, once again suggesting that there is a relationship between subjective culture and usability. However, as the cultural profiles of the test subjects were not identified, this evidence is questionable. The same holds true for the Massey et al. [2001] study.

The lack of assessing the cultural profiles of the test subjects and the negative impact that this has on the results of the studies brings to light that it is necessary to assess the cultural profiles of the test subjects, which is an additional variable that requires inclusion in the conceptual model of usability. Furthermore, as discussed in section 5.2.1, the strengths of the dimensions inherent in each test user should be identified when conducting research into the relationship between culture and usability. This provides additional justification for the need to identify the

cultural profile of the selected users. Consequently, we conclude that the cultural profile of users should be identified when selecting test subjects for a study that investigates the relationship between culture and usability to ensure that the test subjects display the dimensions that are actually being tested.

5.9.2 The Forer and Ford Study

The Forer and Ford [2003] study investigated the effects of heuristics and culture on usability, as discussed in section 2.6.2.2. The primary research methodology used for this study was an experiment, supported by questionnaires and interviews. Three test interfaces were used in this study. The first website had to display superior heuristic usability by conforming to all ten usability heuristics. The second website had to display superior cultural design by conforming to the cultural user interface guidelines, developed by Marcus [2002], that corresponded to the cultural profile of the test subjects. The third website had to accommodate all ten usability heuristics and the corresponding cultural user interface design guidelines. Test tasks were identified in order to compare the user performance achieved on the different interfaces. Following Spool et al.'s [1999] methodology, one task was created for each website with each task containing two informational questions. No judgment questions were asked of users. User testing followed whereby each user completed the three tasks sequentially, providing an accuracy, speed and satisfaction measure for each website. Accuracy was measured in terms of the number of questions answered correctly; speed was determined by the time taken to complete a task; and satisfaction levels were determined by each user completing a Website Satisfaction Questionnaire (WSQ). The WSQ was adapted from Spool et al. [1999]. Informal interviews followed user testing to elicit additional information regarding users' interaction experience.

The following variables were controlled for:

- (1) Identification of the cultural profile of the test subjects – the cultural profile of all test subjects was assessed using Hofstede's [2001] validated and tested Value Survey Model. This confirms the validity of the results in terms of the better performance achieved through the use of an interface that displays a cultural profile matching that of the users. However, as only one cultural profile was used, the results cannot be generalized across different cultural profiles.
- (2) Partial representation – the cultural profile of each user interface component of each test interface was assessed in terms of Marcus's [2002] guidelines. The websites that were chosen as test interfaces were ones where the majority of the components displayed the required cultural dimension. In addition, the test tasks targeted only those components that complied with the requisite cultural dimension side.
- (3) Adaptation of the cultural profile questionnaire – test subjects all had work experience and therefore there was no need to adapt Hofstede's Value Survey Model.

- (4) Compulsory participation – this was avoided by only using test subjects that volunteered to participate in the study.
- (5) The four variables relating to speed of performance were also controlled for. The level of Internet traffic was controlled for by conducting each experiment between the hours of 8.00am and 2.00pm on a Saturday. The same computer equipment was used in each experiment session, thus controlling for differences in hardware platforms. The test tasks all required the same number of navigational decisions to be made, and the same number and length of bodies of text were required to be read to complete the tasks. Consequently user speed of performance was not influenced by differences in these variables.

All the other variables discussed in this chapter were not controlled for, some of which could have distorted the results of the experiment. For example, the relative importance of usability measures was not controlled for, as the findings for accuracy, speed and satisfaction were aggregated to establish overall user performance. Secondly, differences in satisfaction ratings were found between the appearance, metaphors and interaction components of the different test interfaces. However, as the components included characteristics of both sides of a dimension, these differences could have been attributable to differences in the interfaces' cultural profiles rather than a weaker or stronger influence of the components on satisfaction. Thirdly, only some of the performance determinants were controlled for. The test subjects were homogenous in terms of computer literacy, home language, citizenship, and work experience. However, determinants such as cognitive style, ethnic group and verbal fluency were not controlled for, which could have influenced the users' performance.

5.10 Summary

A literature investigation was undertaken to establish whether or not the variables that we thought to have influenced the results of the experiment are valid. This investigation also brought to light additional variables that require inclusion in the conceptual model of usability. The results of our investigation were presented in this chapter, and are summarised in Table 5.11.

Table 5.11 reflects that six categories of variables were identified and investigated. We could not conclusively establish the validity of the four variables relating to subjective culture, the relative impact of components and the nature of the cultural dimension, which presents an opportunity for further research in these areas. However, the theoretical evidence led us to conclude that these variables should at least be considered for inclusion into the conceptual model of usability. All the other variables listed in the table were accepted and will therefore be incorporated into the proposed usability model.

Category	Variable	Validity
Variables relating to Subjective Culture	Cultural Dimension Strengths	To be considered
	Cultural dimension interplays	To be considered
	Relative impact of cultural dimensions on usability	To be considered
	Other subjective cultural dimensions	To be considered
Variables relating to the Interface	Partial representation of cultural dimensions	Accepted
	Usability principles, heuristics and guidelines	Accepted, but include only those that are relevant to the context of use
	Relative impact of components	To be considered
	Nature of the cultural dimension	To be considered
Variables relating to User Acceptance	As listed in Table 5.3	Accepted, but include only those that are relevant to the context of use
Variables relating to Speed of Performance	Hardware platforms	Accepted
	Level of Internet Traffic	Accepted
	Navigational Decisions	Accepted
	Number and length of bodies of text	Accepted
Variables relating to Objective Culture	Nationality and Ethnicity	Accepted
Performance Determinants	Psychological characteristics, job and tasks, knowledge and experience, physical environment and physical characteristics	Accepted

Table 5.11: Validity of Variables Proposed to have Influenced the Experiment

We have also shown that the variables that were accepted as valid, and those that should be considered for inclusion, into the usability model have implications on the way in which experimental research of this nature should be conducted. Specifically, the design will be affected in terms of

- the way in which the usability measures are summated to obtain an overall usability measure;
- avoiding compulsory participation and time limits;
- the way in which pre-existing questionnaires are adapted for use; and
- the order in which test subjects are exposed to test tasks.

In addition, the variables discussed in this chapter also raise doubts as to the validity of the results of four of the studies used in this research. This further confirms that that whilst there is no lack of theoretical underpinnings for cross-cultural usability, there is a lack of explicit demonstration that cultural theories, such as those proposed by Hofstede [2001] are actually transferable [Smith et al., 2004]. Consequently there is a need to determine to what extent

cultural factors actually affect usability and acceptability. This supports the need for further research to be done to establish what the relationship is between subjective culture and usability, thus justifying the need for a more detailed model of usability that will help to establish this more effectively. The conceptual model of usability, encompassing those variables that are known to influence usability and those accepted in this chapter as valid, is presented and discussed in the next chapter.