CHAPTER 2

THEORETICAL FOUNDATION

2.1 Introduction

The influence of culture on system usability is a controversial issue in the field of human-computer interaction. There are many studies in the literature that contend that cultural diversity should be accommodated into the design of user interfaces to enhance usability and performance. However, not everyone in the HCI community seems to agree: The literature contains as many arguments against as supporting the accommodation of culture into user interface design. Those that are in support of culture do not seem to agree on which aspects of culture should be incorporated into user interface design. These opposing views highlight that the problem with cultural issues in the HCI field is that ‘the theoretical foundation for cultural influences on interface design is confused and not yet at the level which would allow designers to predict where culture will have an influence on products, or how these differences can be addressed’ [Dunckley and Smith, 2000, p3].

This chapter presents a review of the literature related to the influence of culture on usability, in an attempt to better understand the issues surrounding this controversy. We begin this chapter by discussing the discipline of human-computer interaction (section 2.2), which forms the context within which this research is based. Definitions relevant to this research and the components of human-computer interaction are presented. This is followed by a discussion of human performance (section 2.3), which is the primary objective of the study of human-computer interaction. Thereafter, we discuss the concept of usability (section 2.4) and related principles as a way of enhancing user performance.

We then turn our attention to the concept of culture (section 2.5). Culture is first discussed in a broad context in terms of an overview of selected metamodels and models, which we show can be simplified into the two categories of subjective culture and objective culture. Arguments for and against the relative importance of culture in HCI are then presented. This is followed by a review of the approaches to culturalization which forms the basis for a metamodel for accommodating both subjective as well as objective cultural variables into the design of interfaces.

Thereafter, we discuss four different models that have been used as a basis for identifying subjective cultural user interface design guidelines (section 2.6). In contrast to the arguments presented in the literature against the use of cultural dimension models, we present theoretical evidence, synthesized from the literature, that such models are applicable to usability. Specifically, we show that Hofstede’s cultural model is applicable to website usability and is therefore a relevant foundation for identifying subjective cultural design guidelines.

We conclude this chapter by noting that there is still confusion regarding the specific aspects of subjective culture that influence usability, and therefore further research is required in order to fully define the relationship between subjective culture and usability.

2.2 Human-Computer Interaction

The term 'human-computer interaction' dates back to the mid-1980s [Kotzé, 2000]. Since then, various definitions for the term have been put forward, including:

- HCI is a 'set of processes, dialogues, and actions through which a human user employs and interacts with a computer' [Baecker and Buxton, 1987, p 40].
- HCI is 'the design of humans and computers interacting to perform work effectively' [Dowell and Long, 1989, p 13].
- HCI is a 'discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them [Hewett, et al., 1992, p 5].
- HCI is concerned with studying and improving the many factors that influence the effectiveness and efficiency of computer use. It combines techniques from psychology, sociology, physiology, engineering, computer science, and linguistics [Johnson, 1997].
- HCI is about 'designing computer systems that support people so that they can carry out their activities productively and safely' [Preece et al., 2002, p 1].
- HCI is 'the study of people, computer technology, and the ways these influence each other' [Dix et al., 1998, p xiii].
- The study of HCI 'seeks to understand what it is about the factors within this total context (of use) that lead to productive and efficient use of computer tools' [Barnard, 1995, p 640].
- HCI is 'a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use', and includes the study of phenomena surrounding interactive systems [Miller, 2003, p 2].

Most of these definitions refer to human performance and usability. For example, Johnson [1999] refers to 'improving the many factors that influence the effectiveness and efficiency of computer use', and Barnard [1995] refers to understanding the factors that lead to productive and efficient computer use. Consequently, we summarise these definitions by viewing human-computer interaction as an interdisciplinary field of study that is concerned with increasing user performance by enhancing the usability of the product.

2.2.1 Interaction Design

The field of human-computer interaction forms part of the larger discipline of interaction design [Preece et al., 2002]. Defined as 'designing interactive products to support people in their everyday and working lives', interaction design aims to develop interactive systems that are easy, effective and enjoyable to use from the user's perspective [Preece et al., 2002], thus enhancing performance [Newman and Lamming, 1995]. As computer technology has evolved over time, successful interaction design requires expertise from a multitude of academic disciplines, design practices and interdisciplinary fields that are concerned with researching and designing computer systems for people. Although HCI is considered to be the best-known interdisciplinary field involved in interaction design, the fields of human factors, cognitive engineering and computer-supported co-operative work are also important for understanding how users react, interact and communicate. These multidisciplinary fields have related academic fields such as ergonomics, informatics, engineering, computer science and psychology, which are also involved in interaction design. In addition, understanding how to design different types of interactive media in effective and aesthetically pleasing ways has also resulted in the need for involvement by practitioners from the design disciplines of graphic, product and industrial design. All of these disciplines are concerned with designing systems to be compatible to users and their goals, but each discipline naturally has a different focus and methodology [Preece et al., 2002].

2.2.2 Components of Human-Computer Interaction

To better understand the concept of human-computer interaction, we present below definitions of the three components of HCI, namely, the human, the computer and the interaction.

2.2.2.1 Human

A human is defined to include different types of users, ranging from children to the elderly, frequent to hesitant users, computer experts and novices [Kotze, 2000]. In the context of HCI, a human who is trying to accomplish something using the technology is referred to as a user. A [human] user can refer to an individual user, a group of users working together, or a sequence of users in an organization, each dealing with some part of the task or process [Dix et al., 1998].

2.2.2.2 Computer

The computer is defined as any technology ranging from the general desktop computer to large-scale computer systems, process control systems or embedded systems [Dix et al. 1998], such as spacecraft cockpits, microwave ovens and electronic games [Hewett et al., 1992]. A computer system consists of people, computers and other components that interact to produce a result that the same components could not produce independently. The system may involve one person working with a particular tool to attain a specific goal or a large group of people working with many tools to meet a shared goal [Bailey, 1996]. A system may or may not include non-computerised components. For example, a manual filing system is non-computerised, whereas a computer-based order-entry system may include computers as well as non-computerised components such as manual business forms and order clerks.

2.2.2.3 Interaction

Any communication between a user and a computer is defined as interaction [Dix et al., 1998], and can be either direct or indirect. Direct interaction involves a dialogue with feedback and control during the execution of a task, in other words, where the user is constantly providing instructions to the system and receiving feedback. The interaction required to use websites is a common example of direct interaction. In contrast, indirect interaction may involve background or batch processing [Kotze, 2000]. The user interface is the means by which the user and the computer interact [Dix et al., 1998]. The user interface requires numerous layers of technology, including [Newman and Lamming, 1995]:

- the application software that supports the user interface;
- the operating system that provides standard services to the user interface and its supporting software;
- system resources such as information storage and printing that are accessed via the user interface and supporting software; and
- the hardware that supports all of the above.

2.3 Human Performance

Computer users and their organizations continuously look for interfaces that will allow them to enhance their performance [Mayhew, 1992]. Frequently, the ultimate goal of systems design is to allow activities to be performed in such a way that more work can be done by a smaller number of people in less time [Bailey, 1996]. The effectiveness of interaction between humans and computers is therefore expressed in terms of performance [Dowell and Long, 1989]. Consequently, the ultimate purpose of an interactive system is to improve the users' performance in their jobs [Mayhew, 1992], and the purpose of the study of HCI is to improve performance, while at the same time providing a safe, comfortable and satisfying experience for the user [Helander et al., 1997]. In order to enhance performance, we first need to understand what it is [Mayhew, 1992].

2.3.1 Measurements of Human Performance

Performance is defined as 'the time and effort expended to complete a task' [Mayhew, 1992, p 30], and is commonly measured in terms of accuracy, speed, satisfaction, and training time [Bailey, 1996].

2.3.1.1 Accuracy

Accuracy relates to the number of errors that are made whilst performing a task. Examples of errors include, typing a 'B' instead of a 'V' [Bailey, 1996] on a QWERTY keyboard, or selecting a menu option that does not contain the required functionality. Many different metrics for measuring task errors have been proposed, including [Bailey, 1996; Nielsen, 1993, Bohmann, 2000]:

- The number of errors made.
- The number of errors made in succession.
- The ratio between errors and successes.
- The proportion of errors made in a given time.
- The number of errors per unit of time.
- The number of users making the same error.
- The number of users completing a task successfully.
- The number of errors made by novice users, or expert users, or the ratio of these errors.
- The number of user errors after a specific time has elapsed.

2.3.1.2 Speed

Speed relates to the rate at which a person works. This may be measured in terms of keystrokes per day, number of invoices captured per hour, or solving a problem before a critical situation develops [Bailey, 1996]. Achieving an acceptable level of speed often means reducing to a minimum the amount of time it takes to complete a task. Alternative metrics for measuring task time include [Bailey, 1996; Nielsen, 1993; Bohmann, 2000]:

- The time users take to complete a specific task or set of tasks.
- The time to complete a task after working with the system for a specific period of time.
- The time it takes a first-time user or repeat user to complete a task.
- The percentage of users completing a task or set of tasks within a specific period of time.
- The percentage of the task completed within a specific period of time.
- The time spent by a user to correct errors.
- The number of tasks of various kinds that can be completed within a specific time,
- The time spent using a help facility.

2.3.1.3 Satisfaction

Constructing systems that allow for increased speed and accuracy does not always guarantee high satisfaction. The system should, at the same time, satisfy the users by being enjoyable, fun, entertaining, helpful, motivating and/or rewarding [Preece et al., 2002]. User satisfaction can be even more important than the performance measures of accuracy and speed for systems that are used on a discretionary basis [Nielsen, 1993; Preece et al., 2002]. The entertainment value of discretionary systems such as games or interactive computing is more important than, for example, speed, as users may want to spend a long time having fun.

Satisfaction is therefore harder to measure as it is a subjective performance measure in contrast to accuracy and speed, which are both objective measures [Shneiderman, 1998; Dix et al., 1998]. Due to the relative difficulty of measuring satisfaction, numerous user satisfaction questionnaires have been put forward in the literature, including the Questionnaire for User Interaction Satisfaction (QUIS), developed by Shneiderman [1998] and refined by Chin, Diehl and Norman [1988], End-User Computing Satisfaction Measures [Doll and Torkzadeh, 1988], and The Post-Study System Usability Questionnaire [Lewis, 1995]. More informal methods have also been used, such as Spool et al.'s [1999] mid-test and post-test questionnaires. User satisfaction has also been viewed in terms of user information satisfaction, for which various instruments and measures have been proposed [Ives, Olson and Baroudi, 1983]. Nielsen [1993] proposes the following simplified metrics for satisfaction:

- The proportion of user statements during task performance that were positive versus critical towards the system.
- The number of times the user expresses clear frustration or clear joy.
- The proportion of users who say that they would prefer using the system over some specified competitor system.

2.3.1.4 Training Time

Training time relates to the amount of time required to develop the skills necessary to perform a new activity [Bailey, 1996]. Also referred to as skill development time, it may be a few minutes or many years. Generally, the less time it takes to train people, the lower the cost of operating the system. As a result, one major goal of systems design is to design activities so that training time is reduced to a minimum and proficiency is maintained after training has been completed.

Bailey [1996] suggests that it is the designers who should be held responsible for reduced human performance, particularly if 'poor design decisions resulted from ignorance of human performance technology'. Thus, in order to improve performance, we also need to know what factors influence it. These are discussed in terms of the human information processing system, cognitive processes and user characteristics.

2.3.2 The Human Information Processing System

The most basic determinant of user performance is the information processing capabilities and limitations of the human mind [Mayhew, 1992]. Human performance takes place through the execution of multiple cognitive processes [Preece et al., 2002; Mayhew, 1992]. These cognitive processes require cognitive resources, which are held in the human information processing system.

Newell and Simon [in Ahituv and Neumann, 1982] developed a formal model of human information processing, which is depicted in Figure 2.1. According to the model, the initial interaction between the user and the environment is triggered by stimuli that attract the human senses (receptors or sensors). The receptors transfer the stimuli to the brain (processor), which uses different types of memory (sensory, short-term, long-term or external memory) to make decisions. The decisions are then routed via the effectors or responders (mouth, leg, hand) back to the environment [Ahituv and Neumann, 1982; Bailey, 1996].

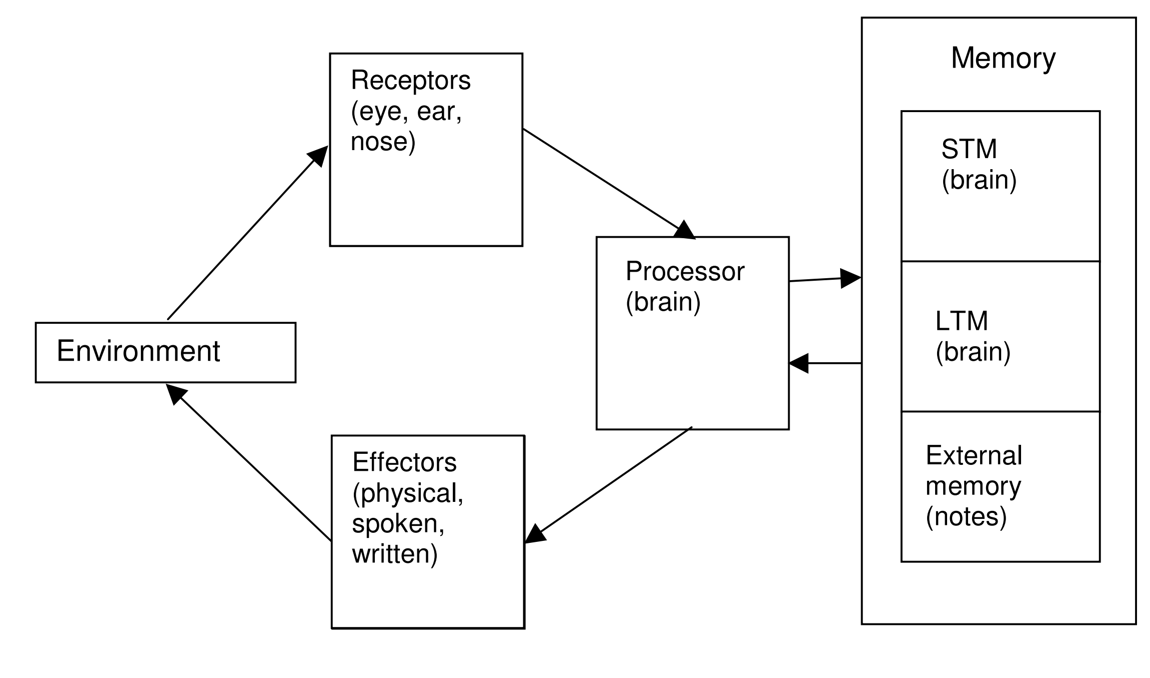

Figure 2.1: Newell and Simon's Model of Human Information Processing [Ahituv and Neumann, 1982]

The receptors, effectors and memory are therefore cognitive resources that are required in order for human information processing, or cognitive tasks, to occur [Mayhew, 1992]. These cognitive resources have inherent limitations and capabilities:

- Short-term memory (STM): STM is used to store information on a temporary basis. It has a limited capacity for holding information – the average person is able to store 7 ± 2 units or chunks [Shneiderman, 1998]. The size of the chunk is dependent on the person's knowledge of the data. For example, a user that is familiar with the stock control (task) domain, and therefore the basic template used for stock codes, would consider the entire stock code as one chunk, whereas a user unfamiliar with the task domain would consider each letter or number of the stock code a separate chunk. As a result, the user's prior knowledge and experience affects the user's ability to remember the stimulus long enough for transfer to long-term memory. Thus, a good knowledge of the task domain increases the number of units available in STM, increasing the user's reasoning and learning skills [Bailey, 1996]. In addition, anxiety reduces the size of memory, as the user is partially absorbed by concerns that are beyond the task at hand [Shneiderman, 1998].

For example, a user may be so concerned about meeting a deadline that he is continuously looking at the time and not concentrating on the task itself.

- Long-term memory (LTM): Long-term memory is more or less a permanent storage facility [Bailey, 1996]; however, access to long-term memory is unreliable, and processing is slow and error-prone [Preece et al., 2002]. Learning and remembering are closely related because the results of learning must be remembered in order for experience to accumulate [Bailey, 1996]. This means that learning and remembering are dependent on long-term memory. Information is more easily transferred into long-term memory if it can be linked to something that is already there. Therefore, new information should be organized or presented in a way that is compatible to what the users are already familiar with to aid in learning and remembering, resulting in enhanced performance. Memory is also affected by the serial position effect, that is, first and last items appearing on a list are recalled with a higher frequency than items in the middle [Barker, 2000; Altman, 2000; Wang et al., 2000].
- Sensors such as eyes and ears also have inherent limitations and capabilities. For example, it is more difficult for the human eye to see blue than other colours [Dix et al., 1998], and the detection threshold for the human eye is a candle flame seen at 30 miles on a clear dark night [Bailey, 1996]. Similarly the detection threshold for the human ear is the tick of a watch under quiet conditions at 20 feet.
- Responders also have basic limitations, in terms of how quickly a hand can move or how accurately it can complete delicate tasks, such as painting or sewing.

2.3.3 Cognitive Processes

In order for performance to take place, a number of cognitive processes are required to be performed. Many of these cognitive processes are interdependent on each other, requiring many of these processes to complete a single activity. Cognitive processes include attention; perception and recognition; memory; learning, reading, speaking and listening; problem solving, planning, reasoning and decision-making; and responding [Ahituv and Neumann, 1982; Preece et al., 2002].

2.3.3.1 Attention

Attention is the process of selecting things to concentrate on, at a specific point in time, from the range of possibilities available [Tesser, 1995]. The most important feature of attention is that it is limited: humans are limited capacity information processors, and our ability to pay attention is very constrained.

From a user interface perspective, the user's attention needs to be attracted from a signal or stimulus that is emitted from the user interface [Preece et al., 2002]. Failure to attract the user's attention would therefore cause the user to '*overlook ... important data because it has failed to* *attract our attention'* [Ahituv and Neumann, 1982, p 13]. Sensing the stimulus requires active participation by the user – '*We can't hear without listening, and we can't see without looking*' [Mayhew, 1992, p 38]. Using computer based information systems this stimulus is usually aimed at the visual or auditory senses [Dix et al., 1998; Preece et al., 2002].

2.3.3.2 Perception and Recognition

When a stimulus attracts the attention of the user, the user then tries to identify it in terms of his or her knowledge base and general context. Perception refers to how information is acquired from the environment and transformed into experiences of objects, events, sounds and tastes [Preece et al., 2002]. Recognition of the stimulus occurs in short term memory, and makes use of context and an existing knowledge base [Mayhew, 1992] that is stored in long-term memory. Thus, perception and recognition involve the other cognitive processes of attention and memory.

People see things differently and behave and respond to the world not as it really is but as they perceive it [CAT, 2000]. Because short-term memory has a very limited storage and processing capacity it is necessary for the users to filter important or relevant signals [Ahituv and Neumann, 1982] from the multitude of simultaneous signals that they are being exposed to. A number of factors can cause users to filter out important data, thus causing faulty perception, and include [Ahituv and Neumann, 1982; Bailey, 1996, CAT, 2000]:

- Context: The situation that people find themselves in.
- Fear or trauma: We avoid seeing things that we do not want to see.
- Internal factors: We are more prone to perceive and recognise stimuli that match our interests, needs, beliefs and expectations than stimuli that oppose them. Preconceptions and culture are both internal factors.
    - Preconceptions result from a user's experiences [Dix et al., 1998]. For example, the financial manager may be expecting a similar monthly sales trend from one year to the next, and may not notice 'just noticeable differences' [Ahituv and Neumann, 1982] in these trends. Consequently, user preconceptions may cause the user to incorrectly interpret data.
    - A person's culture also dictates a person's perceptions; people are conditioned to see and hear in highly constrained ways. For example, Japanese families living in houses with paper-thin walls conditions members of the family to be 'psychologically' deaf [Barker, 2000].
- The nature of the stimulus*:* These include confusion, ambiguity, anchoring and concreteness of data:
    - *Confusion:* People tend to confuse similar auditory signals, for example, F and S, especially in users whose aural sensors are impaired. Similar visual signals may also be confused, for example C and O. This type of confusion may lead to the user not being able to identify the stimulus. For example, a prompt requesting a stock code may

be misread as 'Enter stook oode'.

- Ambiguity: Data can be ambiguous and therefore unrecognisable due to the manner in which it is displayed [Dix et al., 1998]. For example, the letter 'B' can be perceived as the number 13, depending on the context within which it is displayed, the font size and font type used. For example, a stock code 'B24A' may be misread as '1324A'.
- Anchoring*:* A signal may not be identifiable to the user in terms of the user's knowledge base. For example, the use of the signal ●●●● would not be understandable to a user that does not speak Hebrew. Similarly, users that are not familiar with the task domain's jargon would find it difficult to identify stimuli that use such jargon.
- Concreteness of data: Concrete data is data that does not have to be manipulated [Ahituv and Neumann 1982], for example, providing the user with a total Rand value of all the items in the user's shopping cart. The more concrete the data, the more recognisable it is.

2.3.3.3 Planning, Reasoning, Problem Solving and Decision Making

Planning, reasoning, problem solving and decision making are all cognitive processes that involve reflective cognition [Preece et al., 2002], and are also collectively referred to as the analysis process [Ahituv and Neumann, 1982]. They include thinking about what to do, what options are available, working through different scenarios, the consequences of each option, and deciding on which is the best option to follow [Preece et al., 2002]. These processes make extensive use of the processors and memory, as it is during these processes that the stimuli are classified, stored and analysed.

The analysis processes are highly dependent on having accurate perception [Bailey 1996]. Therefore, the processing time required for analysis is partly dependent on the processing time required for perception. For example, data that is difficult to read will take more time to identify, resulting in more time required for the perception process, slowing down the transfer to the analysis processes [Bailey, 1996].

Preconceptions used in deductive and inductive reasoning processes [Dix et al., 1998] may cause errors in the analysis processes. As mentioned earlier, these preconceptions may cause the user to interpret data incorrectly.

The complexity of the problem to be solved often affects the problem solving process. The writing of this document, for example, required the acquisition and analysis of many facts. Conflicting information required additional analysis to solve the problem, and decisions needed to be made on how to structure the document. This complexity required that the problem solving process be broken down into a number of reiterations of the attention, perception and analysis processes.

Reasoning and problem solving require concentration. Any distractions or interferences [Preece et al., 2002] could cause a loss in concentration, resulting in a need to start again, thus reducing speed, and sometimes accuracy, of performance. A number of factors could cause distractions, including

- Slow response time: According to Shneiderman [1998], the system's response time will affect the user's analysis speed, because the user is only able to analyse the stimulus once it has been received. Slow response times also cause the user to become distracted from their task through boredom or irritation, causing loss of concentration.
- Ease of use and learning: The ease of learning and ease of use of the system can also affect concentration. For example, the user may be focused on learning how to use the tool, rather than the problem to be solved, thus causing distractions and even additional anxiety.

2.3.3.4 Responding

Users respond to a computer system most often in the form of a movement, such as pressing a key or moving a mouse to select a menu item. Responding therefore makes use of responders, such as hands and fingers. There are two aspects to responding that influence performance levels: speed of movement and accuracy of movement.

a. Speed of Movement

Movement time affects performance levels, as it affects the time taken to perform the activity required to satisfy the performance objective. According to Fitt' s Lae, the time taken to hit a target is dependent on the size of the target, the distance to be moved, and the relative precision required [Dix et al., 1998]. Other factors such as age, health and environment also affect movement speed. Therefore, the time taken to select a menu item is affected by the menu structure [Newman and Lamming, 1995; Yeates, et al., 1994], as well as the use of either the keyboard or mouse for selection purposes [Dix et al., 1998].

b. Accuracy of Movement

Accuracy of movement is affected by the size of the target and distance that the target has to be moved [Dix et al., 1998]. For example, using Windows Explorer to move a file from one folder to another may result in a higher accuracy of movement when the file is depicted as a large icon, rather than the file name. The distance that the target has to be moved also affects the level of motor skills required by users, and could cause problems for users who are movement impaired, or who lack hand-eye coordination. This could result in faulty movement control [Bailey, 1996].

Accuracy of movement may also be affected by the speed of movement required. It often happens that increased speed reduces accuracy. An example of this can be found in applications that provide a pre-emptive search facility based on the first few letters of a code or word. What we have noticed is that, to prevent the system from inserting an incorrect input after it has completed its search, the user attempts to prevent the search function, by quickly typing in the full code or word. This often leads to input errors as a result of the increased speed of input.

2.3.4 User Characteristics

Although the human information processing system is a general determinant of user performance, in that it is relatively constant across users, there are many other dimensions along which users capabilities can vary [Preece et al., 2002; Dix et al., 1998]. Mayhew [1992] identified a number of such user characteristics that are both internal and external to the user and include the following categories:

- psychological characteristics;
- knowledge and experience;
- job and tasks;
- physical characteristics;
- physical environment; and
- tools.

2.3.4.1 Psychological Characteristics

Psychological characteristics refer to the user's cognitive style, attitude and motivation levels.

- Cognitive style relates to the way in which people approach problem-solving tasks and their ability to solve problems. Different cognitive styles have been identified in the literature, for example inferential/relational and categorical/contextual [Choong and Salvendy, 1999], or verbal/analytic, spatial/intuitive, abstract/concrete, and field-dependent/field-independent [Mayhew, 1992].
- Motivation is any influence that gives rise to performance [Bailey, 1996]. Motivation can be internal or external. Internal motivation is closely related to performance of the work itself, for example, as in the need for competence and self-determination. In contrast, external motivation is related to conditions outside the actual performance of work, such as receiving monetary rewards [Bailey, 1996]. Motivation levels are influenced by negative emotions such as fear, and positive emotions such as interest [Mayhew, 1992].
- Attitude is a mental state which influences an individual's response to all situations [CAT, 2002]. It comprises of knowledge, beliefs and disbeliefs, perceptions, feelings and desires. Factors other than work that influence attitude include age, race, culture and religion and self-efficacy. Self-efficacy is a person's evaluation of his or her ability to perform a task, reach a goal or overcome an obstacle [Baron and Byrne, 2000]. High self-efficacy increases performance because it leads to expectations of doing well, which in turn leads to better performance and a more positive self-evaluation. Self efficacy varies across different kinds of activities. From an HCI perspective, attitude towards

computers and their use can be positive, negative or neutral. Negative attitudes include anxiety about one's ability to understand computers, or fear that computers are going to eliminate one's job [Mayhew, 1992].

Motivation and attitude can affect many of the cognitive processes. For example, users that are not motivated to perform a specific task, or who have negative attitudes towards using computers in general, may not be alert and paying attention to the interface, resulting in important stimuli failing to attract their attention. Low levels of motivation and negative attitudes can also make the users more distractible from the task, thus influencing the analysis processes.

The user's perception of concrete data is dependent on the cognitive style of the user. For example, a user that prefers verbal thinking would consider data displayed in the form of words and equations more concrete than data displayed in pictures, symbols and images [Mayhew 1992]. The cognitive style also affects the amount of data required – for example, a user that prefers to work with detailed information (receptive) would need to identify and interpret more data than a user that prefers to work with summarised data (perceptive). Thus, the user's cognitive style influences the perception process.

2.3.4.2 Knowledge and Experience

This category of performance determinants incorporates system experience, application experience, use of other systems, task experience, typing skills, computer literacy and education, native language and reading level [Mayhew, 1992; Shneiderman, 1998]. These are discussed below.

- System experience refers to knowledge of the particular language or mode of interaction used in a specific system. System experience is also known as syntactic knowledge.
- Application experience refers to whether or not users have had experience with other products of the same genre, for example word processors or spreadsheets.
- Use of other systems is similar to application experience, but relates to the use of other computer-based systems in general, rather than use of other systems of the same genre.
- Task experience refers to knowledge of the task domain, in other words of the general area where the system is being used. For example, to make use of a library system, the user must have knowledge about how libraries work. Task experience is also known as semantic knowledge.
- Typing skill refers to how fast the user can type and is therefore seen as a duplication of keyboard and input skills.
- Computer literacy refers to the user's ability to use the input and output devices common to computer-based systems, as well as their general understanding of computers.
- Education refers to the user's level of education, for example, high school, technical or university degree.

- Native language refers to the user's home language.
- Reading level refers to the grammar and vocabulary that is understandable to the users.

The user's task experience can alleviate the problems associated with ambiguity during the perception process. Using the stock code example, a stock clerk that constantly uses stock codes would know that all stock codes begin with a letter, and not a number. This would make it easier for the stock clerk to realize that the code originally perceived as '1324A' was incorrect, and should be seen as 'B24A'.

However, the user's task, system and application experience could just as easily have a detrimental effect on the perception process. For example, most error prevention messages would attempt to prevent the user from deleting a file by choosing the 'NO' option as the default response to the following message 'You have chosen to delete this file – do you wish to continue?' If the interface designer of the current system has chosen the 'YES' option as the default, the user would, as a result of his preconception, incorrectly identify the stimulus, and push the enter button.

The user's experience with the tools used in the interface can affect the concentration of the user as well. For example, a user that is inexperienced with a particular interaction style, input or output device will be partially focused on learning how to use these tools, rather than on the task at hand. This inexperience can also lead to increased anxiety levels, which in turn reduces concentration and memory capacity. In addition, users become adept at learning how to use new interfaces through practice, which is related to system experience. Therefore, the frequency of use of the interface can lead to better concentration levels.

The user's system and application experience can also influence the response process [Dix et al., 1998]. The language used to communicate with the interface may have an effect on the user knowing how to respond [Dix et al., 1998]. For example, a user who is used to 'erasing' information, may look for an 'ERASE' command on the interface, rather than the more commonly used 'DELETE' command. System and application experience can be influenced by the *frequency of use* of the application. The user's knowledge and experience can also affect the speed of movement attained. For example, a user with good typing skills will be able to input data faster than a novice user, using a keyboard. Similarly, a user that has previous experience with a particular interaction style or application will have a better mental model [Newman and Lamming, 1995] of the system, and will therefore be able to navigate through the interface faster than an inexperienced user.

Movement speed can also be affected by the user's knowledge and experience. For example, a user with good typing skills will be able to input data faster than a novice user, using a keyboard. Similarly, a user that has previous experience with a particular interaction style or application will have a better mental model [Newman and Lamming, 1995] of the system, and will therefore be able to navigate through the interface faster than an inexperienced user.

2.3.4.3 Job and Tasks

This category of performance determinants relates to the user's job and the tasks required to be performed, and incorporates frequency of use, primary training, system use, job category, other tools, turnover rate, task importance and task structure [Mayhew, 1992].

- Frequency of use refers to the number of times and duration of each time that the system is used. For example, users can use the system 5 times a week for 8 hours each time, or once a day for 5 minutes.
- Primary training relates to the way in which the user learns to use the system. This could be through formal training as one extreme, or by self-teaching as the opposite extreme.
- System use refers to whether the use of the system is mandatory or at the discretion of the user.
- Job category refers to the management or operational level that the user falls into within the organizational hierarchy. Typical organizational levels include top, middle and lower management, and operational staff.
- Other tools refer to non-computer based technology such as telephones, calculators and adding machines.
- Turnover rate refers to the staff turnover rate of a given position in an organization.
- Task importance relates to how significant the task is to the user. For example, systems that support appointment scheduling would be very important to the secretary that is responsible for such a task, but not as important as a system that supports the decision-making process for a manager.
- Task structure refers to the repetitiveness and predictability of operations relating to the task.

The frequency of use of an application can alleviate the impact of sensory impairments on the perception process. Frequent use of the application will increase learning and retrieval from long-term memory, allowing the user to remember what the stimulus is, rather than having to try and read or hear it each time [Bailey, 1996; Mayhew, 1992].

2.3.4.4 Physical Characteristics

Mayhew [1992] has identified three dimensions relating to the physical characteristics of the user, namely gender, handedness and colour-blindness.

- Gender relates to whether a person is born as a male or a female.
- Colour-blindness refers to whether a person is able to distinguish between different colours, particularly red and green. Research has shown that 8% of all males are colour-blind [Dix et al., 1998].

- Handedness refers to left- or right-handed users, and those that are ambidextrous.

The user's individual physical characteristics may increase the inherent limitations of the human information processing system. For example, some users may be visually or hearing impaired, leading to difficulties in seeing small icons or hearing low-pitched sounds [Dix et al., 1998]. These sensory limitations may cause the stimulus to fail to attract the user's attention – for example, a colour-blind person may not be able to read an error message displayed in red on a green background.

In addition, a user who is unable to continuously move his arm may find it difficult to use a mouse for very long. The user's physical characteristics could lead to pain and fatigue, which would reduce speed of movement in such a user, thus increasing response time. The user's physical condition may also increase anxiety and reduce concentration. For example, users that are feeling ill or fatigued, or those that are hungry or thirsty, will be concentrating more on themselves than on the job at hand. Consequently, the physical characteristics of the user can also influence performance levels in terms of the analysis process.

Speed of movement also may be affected by the user's physical characteristics. For example, a person confined to a wheelchair who is using an ATM may struggle to reach the necessary keys quickly enough to respond to a prompt for input, thus causing the system to abort before the task is completed.

2.3.4.5 Physical Environment

According to Mayhew [1992], the physical environment comprises lighting, workplace dimension and layout, and furniture and equipment:

- Lighting refers to how well lighted the working environment is, and includes issues such as smoky, dusty conditions that reduce visual acuity.
- Workplace dimensions and layout refer to the amount of space available within the workstation as well as how easily accessible the equipment and furniture are to the user.
- Furniture and equipment relate to the furniture and equipment used by the user in performing one or more tasks.

The physical environment plays an important role in the ability of the stimulus to attract the user's attention. For example, a noisy working area would make it difficult for the user to hear auditory stimuli, whereas poor lighting or smoky conditions could make it difficult to see visual signals. These problems would be exacerbated for users who are already visually or aurally impaired. In addition, an order clerk using an order entry system could be exposed to visual, auditory and haptic stimuli from the computer interface, the customer and the environment. Because sensory memory is limited, the stimulus may be erased by new inputs being received simultaneously [Bailey 1996], thereby causing the user to overlook the stimulus.

.

2.3.4.6 Tools

Mayhew [1992] refers to the user's tools as the means of conducting tasks, such as spreadsheets to perform calculations and word-processors to compile written documents. Although the other six determinants can be manipulated to improve performance, the user's tools are the only determinants that are under the designer's control. However, 'if the designer knows the user's level on all other performance determinants, the designer will be able to design the tool to more effectively fit the user' [Mayhew, 1992, p 72]. System designers must therefore take into consideration the user's job and tasks, knowledge and skills, motivation and attitude, physical characteristics, and information processing capabilities when designing the user interface.

2.4 Usability

Making products and systems easier to use will reduce the cognitive load on the users, thereby enhancing user performance. Using computers to complete tasks always results in needing to perform at least one functional as well as one operational task simultaneously [Mayhew, 1992]. Functional tasks relate to the content of the problem at hand, whilst operational tasks refer to the means of solving the problem. Consequently, these two tasks will compete for cognitive resources, thus resulting in degraded performance of one or both of the tasks, and of the interaction overall. Therefore, the most obvious way of enhancing performance is to design systems that will reduce the amount of cognitive resources required to complete both operational as well as functional tasks [Mayhew, 1992; Baecker and Buxton, 1987].

Solving a problem by using computer software that is difficult to understand, hard to learn and complicated to use will result in an increase in the cognitive resources required, thus reducing performance even further. Consequently, HCI is ultimately about making products and systems easier to use, and matching them more closely to user needs and requirements. This notion has been termed 'usability' [Miller, 2003], and is tested in terms of the resultant performance achieved from using a computer to complete tasks [Nielsen, 1993].

2.4.1 Definitions

Usability is defined as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [ISO 9241, 1997; CHI-SA, 2000; Dix et al., 1998], where

- effectiveness is defined as 'the accuracy and completeness with which the goals of using the system are achieved';
- efficiency is defined as 'the amount of resources, including time, money and mental effort, that is expended to achieve the intended goals'; and

- satisfaction is seen as describing the comfort and acceptability of the work system to its users and other people affected by its use [Dix et al, 1998, p 192].

Usability has remained a fuzzy concept that has been difficult to measure and evaluate [Bevan, 1995]. To make usability a more measurable concept, a number of usability characteristics have been identified [Dix et al., 1998; Preece et al., 2002; Nielsen, 1993; Mayhew, 1992; Shneiderman, 1998]. Hence, usability is not a one-dimensional concept, but rather consists of multiple characteristics that should be considered when specifying and measuring the usability of products [Nielsen, 1993; Preece et al., 2002].

However, not all of the usability goals and related metrics are relevant to every product [Preece et al., 2002]. A system is used within a specific context, in other words, by users with specific characteristics, to perform specific tasks in a specific environment. Thus, usability is a function of the context of use [Bevan, 1995], and that context can change the usability of the system. For example, Nielsen [1993] points out that satisfaction can be a very important measure for systems that are used on a discretionary basis, such as games or home computing. For such systems, their entertainment value is more important than speed as people may want to spend a lot of time having fun. Consequently, the ideal way to measure usability is to identify the characteristics that make the system usable, and then measure whether these characteristics are present in the system. The relative importance of these characteristics is therefore dependent on the context of use, in other words, the environment, the task at hand and who the intended users are [Preece et al., 2002]. As a result, usability characteristics and related metrics should be chosen that are appropriate to the context of use.

2.4.2 Context of Use

The context of use identifies who the intended users of the product are, what kind of tasks the users will perform with the software, and the equipment and environment that will be used in conjunction with the software [Kirakowski and Cierlik, 1999]. A number of different sets of contexts of use have been put forward [Bevan, 1995, Kirakowski and Cierlik, 1999; Maguire, 1998]. Two of these sets are illustrated in Table 2.1.

Kirakowski and Cierlik [1999] identify three contexts, namely the user, the task and the environment. Bevan [1995] identifies four contexts, namely the user, the task, the environment and the equipment, where the environment context is further divided into organizational, physical and technical categories. Kirakowski and Cierlik's environment context is divided into the four categories of social, organization, technical and physical environments, as opposed to Bevan's three (not including the social context of Kirakowski and Cierlik). It appears, however, that some of the variables identified by Bevan in the equipment context have been duplicated in the technical category of the environment context, and we will therefore treat the two as a single category. Kirakowski and Cierlik include categories in the user context for the user's job characteristics, attitude and motivation, and role, while Bevan identifies motivation and attitude as variables relating to the user's personal traits, and the user's job characteristics are identified as separate variables in the task context and organization category of the environment context. These contexts of use will be described in more detail in sections 2.4.2.1 to 2.4.2.3. .

Context	Bevan [1995] Category	Kirakowski and Cierlik [1999] Category
User	Skills and Experience	Skills and Knowledge
	Personal Traits	Physical Attributes
		Attitude and Motivation
		Job Characteristics
		Role
Task		Task Execution
		Task Flow
		Task Demand on Users
		Task Safety
Equipment	General	(Addressed as part of Environment)
	Specifications	
Environment	Organisational	Organisational
	Technical	Technical
	Physical	Physical
		Social

Table 2.1: Comparison of Contexts of Use

Many of the variables identified in these contexts of use have also been identified as characteristics that determine user performance [Mayhew, 1992], which were discussed in detail in section 2.3.4. Given that system usability is tested in terms of the resultant performance achieved from using the system, this duplication is not surprising. We will defer further and more detailed discussions of these duplications to Chapter 6, where we present a revised and expanded context of use based on the findings of this research.

2.4.2.1 The User Context

The user context takes into consideration all the characteristics of the user that may influence his or her ability to use the interface. A product can be usable for a user with specific characteristics, but unusable for users with differing characteristics: 'Users share common capabilities but are individuals with differences, which cannot be ignored' [Dix et al., 1998]. Thus, the aim of analyzing the user context of use is to identify the characteristics of the intended users that may influence the usability of the product [Maguire, 1998].

Potentially relevant user characteristics, as synthesized from the contexts of use proposed by Bevan [1995], Maguire [1998] and Kirakowski and Cierlik [1999], as well as the performance determinants identified by Mayhew [1992] are:

- Skills and knowledge, including:
  - Product experience: hands-on experience with a particular software application, such as MS Word or Pastel Accounting.

- System knowledge / general computer experience: knowledge about computer-based systems in general. For example, a user may have basic computer literacy knowledge or may have experience with similar products but no experience with a particular product.
- Task experience: the amount of training or experience that the user has had in the task domain.
- Organisational experience: the user's prior work experience with specific organizations or organisations in general.
- Training: the type of, and manner in which, users have learned to use computers, specific software products, or perform their job functions.
- Keyboard and input skills: the user's skills in using input devices, such as the keyboard, mouse or joystick.
- Qualifications: the user's level of education that he/she has achieved, for example, Grade 12, degree, diploma or postgraduate studies.
- Linguistic ability is defined as the user's ability to speak, read and understand a particular language.
- General knowledge is defined as the user's general knowledge of the world, for example, technology, politics, economics, agriculture and sociology.

* Physical attributes, including:
    - Age: the specific age of the user or the age group to which the user belongs, for example, between the ages of 18 - 22.
    - Gender: whether the user is a male or female.
    - Physical capabilities: the user's levels of mobility, speed and strength, vision and hearing, and any associated disabilities or limitations.
* Psychological attributes, including:
    - Attitude: the attitude of the users to their jobs and tasks, to a specific software product, to information technology in general, and to the organization that employs the users.
    - Motivation: the level of motivation to use a computer-based product.
    - Intellectual abilities: the user's ability to process information, make decisions and solve problems, that is, the user's cognitive abilities.
* User role, indicating whether the user is:
    - a direct user who operates the system directly,
    - an indirect user, who is indirectly affected by someone else's use of the software, for example, a user having to handle outputs or inputs to the software,
    - supporting user, who is involved in supporting other users making use of the system, such as help desk staff, or
    - monitoring user, where the user employs the system to monitor or check the work of other users of the system.

.2.4.2.2 The Task Context

The task context considers all characteristics of the task that could influence the user while using the computer to perform that task. We also include the variables listed under Kirakowski and Cierlik's job characteristics category within this context of use, rather than in the user context of use, as jobs comprise of a multitude of tasks. For the purposes of analyzing the task context, a task is 'considered to be an activity which starts from an initial known state of the human, the computer or the environment, and continues until it reaches a known end state' [Kirakowski and Cierlik, 1999, p 1]. Although task analysis is a necessary technique for identifying functional requirements, for the purposes of identifying the task context it is recommended that one should rather concentrate on defining the main tasks that are relevant to each class of user using the application.

Potentially relevant variables relating to the task context as synthesized from the contexts of use proposed by Bevan [1995], Maguire [1998] and Kirakowski and Cierlik [1999], are:

- Task characteristics, including:
  - Task goal: the aim or objective of the task, for example, taking orders from customers or reconciling the cashbook.
  - Task breakdown: the activities that need to be performed in order to complete a specific task.
  - Output of the task: the information that the task produces.
  - Risk resulting from error: the precision required in the completion of the task in order for the output to be of use or not to cause dangerous situations. This includes the amount of monetary loss that may arise as a result of an error in the task. Very often the task goal will directly influence the amount of risk.
- Task execution, including:
  - System use: whether or not the use of a system to complete the task is mandatory or voluntary, in other words the organizational policy on the use of systems for specific tasks.
  - Criticality of task output: how important it is that the task is carried out or completed. This is particularly relevant for safety-critical and mission-critical tasks.
  - Task autonomy and discretion: the extent to which the task is carried out by the user without supervision or quality checks.
  - Frequency of the task: how often the task is performed, for example, daily or monthly, or once or twice in a lifetime.
  - Duration of task: the time spent performing the task, for example, minutes, days or months.
  - Frequency of events: how often events take place that require response in terms of executing a task. The frequency of events is therefore seen as relating to the frequency of the task, as the more often the event occurs, the more often the task will be performed.

- Flexibility of the task: whether or not the task must be completed in a certain way, in other words, whether all the activities need to be performed, whether all the activities need to be performed in the same order, and whether all the activities need to be performed using the system.
- Task postponement: whether or not the user can postpone the completion at any stage, and if so, the kind of information that needs to be maintained during the postponement period.

- Task flow, including:
  - Task side effects: the extent to which any side effects or by-products of the task are desirable or unwanted.
  - Dependencies: information, events or other resources that are required before the task can be executed or completed. For example, having a credit card or receiving information about current exchange rates.
  - Linked tasks: whether or not the task is carried out concurrently with other tasks
  - Task pacing: whether any of the activities need to be performed within a specific time constraint.
- Task demand on users, including:
  - Physical demands: the physical capabilities that are required by the user in order to perform the task. This variable correlates indirectly to the physical capabilities variable defined under the user context of use.
  - Mental demands: the intellectual capabilities that are required by the user in order to perform the task. This variable correlates indirectly to the intellectual capabilities variable defined under the user context of use.
- Task safety, including:
  - Physical / psychological safety: the implications of the task on the physical or psychological well-being of the users (direct or indirect).
  - Information safety: the implications of the task on the integrity of the user's information environment (for example, data stored on the user's hard drive) or a wider information environment (for example, data stored on a remote server).

2.4.2.3 The Environment and Equipment Context

The environment is an important context to consider as it sometimes happens that a certain category of user carries out exactly the same task in two different environments [Kirakowski and Cierlik, 1999]. There are many aspects of the environmental context, which have been grouped into various categories, including organisational, social, physical and technical. Potentially relevant characteristics relating to each of these environments as synthesized from the contexts of use proposed by Bevan [1995], Maguire [1998] and Kirakowski and Cierlik [1999], are:

- The organisational environment, including:
  - Organisational structure: management and communication structures that affect the way in which the user works or uses the system.

- Organisational culture: general organizational policies and procedures that may influence the users' work practices, as well as any policies and procedures specific to the way in which the user interacts with computer systems.
- Performance monitoring: the extent to which users' work is monitored and the effect that such monitoring has on the user.
- Performance feedback: whether or not the organization provides feedback to the users in terms of their monitored performance.
- Hours of work: the number of hours that an employee is expected to work per day, as well as anything special about the hours of work required, such as long or anti-social hours, shift work or flexitime.
- Work autonomy: whether the user works on demand by circumstances or other people, or whether the users can choose where and when they carry out their jobs.

- The social environment, including:
  - Assistance available: the extent to which human assistance, for example, help desk support, is available to the user, and the extent to which obtaining this assistance interferes with the execution of the task at hand.
  - Interruptions: the extent to which the working environment is intrinsically disruptive, which is measured in terms of the number of interruptions as well as the duration of the interruption.
  - Use of single or multi-user environments: the extent to which users interact with other users in the environment, which may or may not be task related, for example, co-workers using the same or similar systems to carry out different tasks.
- The physical environment, including:
  - Workplace conditions, which includes issues such visual and auditory conditions;
    - ♣ Location type: the type of setting, such as a standard office, laboratory or training classroom, home or other informal setting, or kiosk, such as an ATM.
    - ♣ Atmospheric, auditory, thermal and visual conditions: the extent to which the setting is affected by unusual noise, temperature, vibrations or light conditions, and the amount of fluctuations experienced in these conditions.
  - Workplace safety, which includes
    - ♣ Health hazards: any intrinsic hazards to the users' health that should be guarded against, for example, danger from pickpockets in a kiosk setting or inflammable chemicals in a laboratory setting.
    - ♣ Protective clothing and equipment: the extent to which protective clothing and/or the use of protective equipment may affect the users' ability to see displays or manipulate input devices.
  - Workplace design, which incorporates variables relating to the amount of space allocated to users, the furniture and location of the workplace.
    - ♣ Furniture: furniture or other equipment required for the successful operation of the system or completion of the task.

    - ♣ Space: the amount of space available that may constrict the user while executing a task.
    - ♣ Posture: the extent to which users' will be required to habitually adopt an unusual or uncomfortable posture while using the system or any other equipment or material required during the execution of the task.

- The Technical Environment, which includes:
    - Standalone / networked: whether the product will be used on a standalone computer or on a network.
    - Software: the software environment in which the system will run, for example the operating system. This includes software that is required to be used in conjunction with the system, for example web browsers or email applications.
    - Hardware: the hardware that is required in order to run the system, in other words, the hardware that the user will interact with.
    - Other materials and equipment: computerised and non-computerised equipment, such as telephones, calculators and overhead projectors that may need to be used in conjunction with the system. This includes reference materials such as user manuals or procedure manuals.
    - Main application area: the functional area of the organization that is supported by the technology, for example, human resources or manufacturing.
    - Major functions: the specific business functions or activities that are supported by the technology, for example, payroll or order-entry.

2.4.3 Usability Characteristics

Usability characteristics are often referred to as goals, principles, heuristics or guidelines. The use of these terms is confusing, as they are often used interchangeably and in different combinations. For example, some of the literature refers to 'guidelines' as the most general of the terms, as they are seen to relate to all forms of guidance, thus incorporating heuristics, principles, and goals [Preece et al., 2002]. In contrast, other literature sources [Newman and Lamming, 1995; Nielsen, 1993] use the term 'guidelines' to describe the way in which usability principles are instantiated; in other words, guidelines are seen to provide the 'how', as opposed to the principles that state the 'what'. For the purposes of this section, the following definitions of these terms, which have been synthesized from the works of Dix et al. [1998], Mayhew [1992], Nielsen [1993], Preece et al. [2002] and Newman and Lamming [1995], will be used:

- Usability goals: the high-level usability objectives of the system, for example, the system should be efficient, effective and safe to use, and must be easy to learn and remember [Mayhew, 1992].
- Usability metrics: are identified for each of the usability goals to further enhance the measurability of usability. These metrics allow each usability goal to be measured in terms of its contribution to the three categories of usability (effectiveness, efficiency and

satisfaction) [Dix et al., 1998]. Given that the definitions of usability and performance are similar, metrics therefore measure how each usability goal can improve user performance [Preece et al., 2002].

- User experience goals: what the interaction of the system feels like to the user, for example, how satisfying, fun, motivating and rewarding the system is for the user [Mayhew, 1992; Preece et al., 2002].
- Principles: These are generic guidelines that are applicable to all user interfaces. They encompass the overriding and fundamental requirements of making an interface user-friendly [Nielsen, 1993], and are based on cognitive psychology theories [Mayhew, 1992]. Examples of principles proposed in the literature include Shneiderman's [1998] eight golden rules, Mayhew's [1992] general principles of user interface design, IBM's ease of use principles [IBM, 2004], and Dix et al.'s [1998] general usability principles. These are discussed in more detail in section 2.4.5.
- Heuristics: are closely related to principles. When principles are used in practice, they are referred to as heuristics [Preece et al., 2002]. Examples of heuristics include Nielsen's [1993] 10 usability heuristics and Instone's [1998] adaptation for web usability. Heuristics are discussed in more detail in Chapter 5.
- Guidelines: provide general advice about the usability characteristics of the interface. They instantiate the principles, in other words, guidelines provide the 'how', as opposed to the principles which state the 'what' [Nielsen, 1993; Newman and Lamming, 1995]. Guidelines can be category-specific or product-specific. Category-specific guidelines are relevant to specific types of systems such as accounting packages or word processors, or graphical user interfaces or DOS-based systems. Ensuring that an undo button is visible on the screen is an example of a category-specific guideline for graphical user interfaces. Product-specific guidelines are relevant to an individual product, such as MS Word or MS Windows. An example of a product-specific guideline for MS Windows is that different icons must be used to represent different types of files [Nielsen, 1993].
- Rules and standards: specify how the interface should look to the user. Rules are generally developed by vendors or corporations, such as Microsoft, IBM and Macintosh. As a result, they become internationally recognized as international standards [Nielsen, 1993]. An example of a Microsoft standard would be the use of a book icon to represent help files.
- Usability characteristics: we propose this term as a way of referring to all forms of guidance, encompassing goals, principles, heuristics, guidelines and standards.

A detailed discussion and comparison of all of the usability characteristics is outside the scope of this dissertation. We do, however, discuss some of the usability goals and principles that are most relevant to this research in more detail below. A detailed discussion of heuristics and guidelines is deferred to Chapter 5.

.

2.4.4 Usability Goals

Usability goals are based on the concept of reducing the cognitive resources required when using information technology as a means of performing tasks [Preece et al., 2002]. Consequently, measurement studies are important for assessing whether the usability goals have been met [Nielsen, 1993]. Such measurement studies require usability metrics. For each usability goal of interest, several different levels of performance can be specified. Nielsen [1993] contends that at least a minimum performance level should be set for each goal, but that a more detailed metric can also include a planned level that one is being aimed for as well as the current level of performance. Therefore, metrics are specific objectives that can allow usability goals to be assessed in terms of how each one can improve user performance [Preece et al., 2002]. We discuss below some of the usability goals and metrics in terms of their relation to user performance measures

2.4.4.1 Learnability

Learnability relates to how easy the system is to learn how to use [Preece et al., 2002]. Users normally do not take the time to learn the entire interface before they start using it, instead, they often start using a system as soon as they have learned part of the interface [Nielsen, 1993]. A simple way to measure initial ease of learning is to select users who have never used the system before, and measure how long it takes them to reach a specified level of proficiency. The most common way of measuring this proficiency is to identify a specific task or set of tasks that the user should be able to complete successfully, and then to measure how long it takes for the user to reach that stage [Nielsen, 1993]. Consequently, learnability relates directly to the user performance measure of training time. However, it is evident that learnability can also be measured in terms of the accuracy and speed with which users can complete tasks after a specified period of time. For example, the number of errors made after a specified amount of training time has elapsed should be less for a system that is easy to learn how to use. Therefore, accuracy metrics can also be used to indicate learnability, such as

- the number of users making the same error;
- the number of users completing a task successfully; and
- the number of errors made by novice users, or expert users, or the ratio of these errors.

Similarly, the amount of time taken to complete a task will also indicate the ease of learnability of the system. Measures of speed can also be used to indicate learnability, such as

- the average time users take to complete a specific task or set of tasks;
- the percentage of users completing a task within a specific period of time; and
- the average number of tasks completed within a specific period of time.

As these metrics were identified for accuracy and speed (see sections 2.3.1.1 and 2.3.1.2), learnability is also related to the performance measures of accuracy and speed.

2.4.4.2 Memorability

Memorability refers to how easy it is to remember how to use the system, once it has been learned [Preece et al., 2002]. Essentially, the goal is to design the system in such a way that the casual user can return to the system after some period of not having used it, without having to learn everything all over again [Nielsen, 1993]. According to Nielsen [1993], the best way to measure memorability is to measure the performance of casual users that have not used the system for some time. This method of measurement once again suggests that memorability, like learnability, is related to the performance measures of accuracy and speed.

2.4.4.3 Efficiency

Efficiency refers to the extent to which a system supports users in carrying out their tasks [Preece et al., 2002], so that once the user has learned the system a high level of productivity is possible [Nielsen, 1993]. Preece et al. [2002] suggest that an efficient way of supporting common tasks is to design the system in such a way as to reduce the number of buttons or key presses used. Therefore, efficiency can be expressed in terms of the time taken by a user to complete a task, thus relating to speed of performance [Nielsen, 1993]. The time spent in correcting errors will also influence overall task time, and therefore efficiency can also be related to the performance measure of accuracy.

2.4.4.4 Safety

Safety relates to two aspects of using the system: (1) protecting the user from dangerous conditions and (2) undesirable situations [Preece et al., 2002]. Dangerous conditions relate to the user's physical environment, such as using X-ray machines that emit harmful rays or working in a nuclear plant where users can be exposed to dangerous chemicals. Undesirable situations refer to helping users avoid the dangers of making catastrophic errors [Nielsen, 1993]. Safety in terms of error prevention is again related to the performance measure of accuracy. The effects of errors on the time taken to successfully complete a task also relate the usability goal of safety to performance speed [Nielsen, 1993].

2.4.4.5 Satisfaction

Systems must be pleasant to use so that users are subjectively satisfied when using it, in other words, the users must like the system [Nielsen, 1993]. In contrast, Preece et al. [2002] suggest that satisfaction is not a usability goal, but rather a user experience goal. Usability goals are related to improving efficiency and productivity, while the user experience goals relate to user satisfaction, and include factors such as fun, enjoyment, entertainment and helpfulness. In addition, they point out that usability goals and user experience goals are often conflicting and require tradeoffs, for example, it would be difficult and unnecessary to design a process control system that was both fun and safe. Irrespective of whether satisfaction is seen as a usability goal or a user experience goal, it is directly related to the user performance measure of satisfaction.

2.4.4.6 Utility

Nielsen [1993] contends that usability and utility are separate concepts, in that utility is concerned with whether the system can do what is needed, whilst usability is how well the users can use that functionality. In contrast, Preece et al. [2002] view utility as a usability goal, in that it refers to the extent to which the system provides the right kind of functionality to allow users to perform required tasks, in the way that they want to or need to. We suggest that these two definitions are similar, in that they both take into consideration the user needs when performing tasks. Preece et al. [2002] point out that a system may force a user to perform a task in a way that is inefficient or unknown to the user, resulting in accuracy, speed and satisfaction being impeded. Therefore, utility can be related to all three measures of performance.

2.4.5 Usability Principles

Of the sets of principles listed in section 2.4.3, the set proposed by Dix et al. [1998] seems to be the most comprehensive, as they are general enough to be applied to the design of most interfaces. These principles are organized into three distinct categories, namely learnability, flexibility and robustness.

(1) Flexibility considers the different ways that the user and the system exchange information. This category includes the principles of dialogue initiative, multi-threading, task migratability, substitutivity and customizability.

(2) Robustness relates to the level of support provided by the system to the user in determining successful achievement and assessment of goals. This category includes the principles of observability, recoverability, responsiveness and task conformance.

(3) Learnability relates directly to the usability goal of the same name, and considers the ease with which new users can begin to effectively interact with the system and to achieve maximal performance. Principles included in this category are predictability, synthesizability, familiarity, generalisability and consistency. Consistency is one of the fundamental principles of usability [Dix et al., 1998] that is often used to improve human performance, irrespective of the context of use. However, we propose that compatibility is more appropriate for enhancing performance. Consistency and compatibility are discussed in more detail in sections 2.4.5.1 and 2.4.5.2.

Like usability goals, usability principles are based on the theory of cognitive psychology, a field of study that attempts to understand the general strengths and weaknesses of the human mind [Mayhew, 1992]. These principles correlate to the usability goals, for example, task conformance is related to utility, and all the principles in the learnability category are naturally related to the usability goal of learnability. Consequently, the principles that are relevant to a particular product are dependent on the context of use.

2.4.5.1 Consistency

In its narrowest sense, consistency is defined as similarities within a product [Mayhew, 1992], for example, the same information should be located in the same place and displayed in the same way on all screens [Nielsen, 1993]. Also known as internal consistency [Grudin, 1989], this allows the user to become familiar with the system more easily, thus reducing the cognitive load because it reduces the amount of learning required by the user.

Other forms of consistency have also been proposed such as external consistency and consistency to the real world [Grudin, 1989]. External consistency, also referred to as product compatibility [Mayhew, 1992] relates to the design of interfaces being consistent to the design of other interfaces that users are familiar with, for example, using the same menu options in a word-processing and spreadsheet application. External consistency is seen to assist in the transfer of skills, as the user is able to reuse his/her knowledge gained from prior experience with one application, on a new application. Consistency to the real world relates to using metaphors, terminology and other concepts that are known to the user, and is also often referred to as familiarity [Mayhew, 1992]. Thus, external consistency and consistency to the real world are both seen as ways of reducing the cognitive load required during the initial process of learning how to use the system.

However, Grudin (1989) argues that consistency in interface design can be harmful to the user, rather than helpful. He points out that interface objects should rather be compatible to the user's tasks, physical and psychological capabilities and limitations. For example, although the optimum keyboard layout is not consistent to the real world, and therefore does not support ease of learning, better performance is achieved using this layout because it is compatible with the factors that govern typing performance, such as the balance of workload across hands and fingers and the average length of finger trajectory. As another example, Grudin (1989) shows that consistency in menu defaults reduces performance, as the default menu option is normally the option that is most likely to be selected. Menus are used for a variety of tasks and therefore the option that is most likely to be selected will depend on the task being performed. This suggests that the best default option should be chosen based on the task being performed, rather than on the basis of maintaining consistency to previous interactions.

The essence of Grudin's (1989) argument is that design decisions should be based on the user, their tasks and their environment. This essentially refers to the need to design systems that are compatible to the context within which the system is being used.

2.4.5.2 Compatibility

According to Mayhew [1992], compatibility refers to the ability of the interface to be consistent to the user's profile. Many of the principles proposed in the literature can be seen to relate to compatibility, including the three types of consistency discussed above. For example, the following principles proposed by Mayhew [1992] can be related to compatibility:

- Task and workflow compatibility suggests that the design should be compatible to the user's job and tasks profile.
- Flexibility relates to the ability of the interface to accommodate variations in the skills and preferences of users. Although Dix et al. [1998] view flexibility as a category containing many principles, Mayhew's [1992] definition is narrower in that flexibility is viewed as relating to the ability of the user to control and adjust the interface. This suggests that if the interface is flexible, it can be adjusted to be compatible to the user's knowledge and experience, psychological and physical characteristics, and their physical environment.
- Easy to use and ease of learning means that the interface should accommodate novice as well as expert users. This suggests once again that the interface should be compatible to the user's knowledge and experience profile.
- Direct manipulation makes use of metaphors, particularly when designing icons, can also relate to compatibility. For example, a user would understand a 'save' icon if it was represented as an open file, as the icon would be compatible with a 'real world' experience of filing a document in a physical file.

The principles proposed by Dix et al. [1998] and listed above can also be seen to support the notion of user compatibility. For example:

- Predictability requires that an interface should be consistent in its internal behaviour, as well as compatible to the user's mental model. This means that the interface should be compatible with whatever created the user's mental model of the system, such as previous experiences with other or similar systems, task domain knowledge, and the tasks to be performed.
- Generalizability requires that users must be able to apply their knowledge on how to achieve a particular goal, to another situation where the goal is in some way similar. This once again suggests that the system must be compatible to the user's prior knowledge and experience.
- Dialogue initiative refers to whether the system or the user is in control of the interaction. With system pre-emptive dialogue, the system is in control of the interaction. This reduces flexibility in terms of user control, and could therefore reduce compatibility to the user's job and tasks, particularly if the system forces the user to perform a task differently to the way the user is used to. User pre-emptive dialogue can also reduce compatibility to the user's prior experience, as the user may not be experienced enough in the task domain or with the system to effectively control the dialogue.
- Multithreading relates to the ability of the interface to be compatible to the tasks and habits of the user, for example, having to swap between tasks.
- Task migratability relates to the ability of the user to pass control over to the system, or maintain some level of control over the task being performed. A user may choose to

pass control over to the system due to inexperience or anxiety, thus task migratability increases compatibility with the user's knowledge and experience and psychological characteristics.

- Substitutivity refers to the ability of the system to cater for equivalent values of input and output, for example, substituting inches for centimetres. Substitutivity therefore enhances compatibility to the user's real world and task domain.
- Customisability refers to whether the interface can be modified to suit user preferences and skills. There are two types of customizability: Adaptability, which is user-initiated, allows users to modify the system according to their preferences and skills. Adaptivity, which is system-initiated customizability, can reduce the effects of system pre-emptive dialogue on compatibility to user skills, by automatically adjusting the interface to accommodate novice and expert users. However, adaptivity can also decrease consistency in the interface's behaviour. For example, if the system classifies a user as an expert one day, and then as a novice the next day, the system may disable menu options that were previously available to the user when the system classified that user as an expert.
- Responsiveness refers to the amount of time that the user is given to respond to the system, as well as the time that the system takes to respond to a user command. Responsiveness can reduce the interface's compatibility to the user's physical and psychological characteristics. For example, a system that requires quick responses would not be compatible to users with physical impairments, or lack of experience that would slow down the user's response.
- Task conformance refers to the ability of the system to support all of the tasks that the user wishes to perform (task completeness), in the way that the user understands them and wants to perform them (task adequacy). Hence, task conformance relates to the user's job and tasks, as well as the user's preferred way of working.
- Recoverability refers to the ability of the user to detect, diagnose and recover from errors. If error messages and recovery actions are compatible to the user's prior knowledge, users will be better able to detect and deal appropriately with errors.
- Observability refers to the ability of the user to evaluate the internal state of the system from its perceivable representation. Representations such as messages and icons should be compatible to the user's prior knowledge and visual abilities in order for the user to be able to notice and interpret them.

In addition, the three types of consistency discussed above can also be viewed in terms of user compatibility.

- Familiarity refers to consistency in concepts, terminology and spatial arrangements of the real world, and therefore suggests that the interface should be compatible to the user's prior knowledge and experience of the real world.
- Product compatibility suggests that the interface design should be consistent with other

products or applications, and therefore should be compatible with the user's application and system experience profile.

- Internal consistency requires that the interface design be compatible to the user's knowledge and experience with other parts of the same application.

The above analysis of the usability principles has highlighted that compatibility can be viewed in terms of the performance determinants of the user's job and tasks, knowledge and experience, physical characteristics, psychological characteristics, physical environment and tools. This leads us to believe that an interface that is compatible to these determinants is required in order to achieve usability. Furthermore, interfaces need to be customisable to cater for the ever-changing characteristics of the user. These determinants also correlate to the four contexts of use (user, task, environment and equipment), which provides further support for the notion that usability is obtained through compatibility to the context of use.

As we have also shown that consistency in some cases can be harmful, rather than helpful, to user performance, we propose that compatibility is a more appropriate usability characteristic than consistency for improving performance. What is needed therefore is an approach that will help designers to design products from the user perspective. User Centred Design (UCD) is such an approach, consisting of the techniques, methods, processes and procedures for designing usable systems [Rubin, 1994, Dix et al, 1998, Preece, et al., 2002]. Even more importantly, UCD is based on the philosophy that places the users, their environment and their tasks at the centre of the process, thus ensuring compatibility of the design to the contexts of use [Rubin, 1994].

2.5 Culture

Bevan and Macleod [1994] point out that 'individual differences between users such as those due to cultural differences and prejudices may influence the use of a product in the real world' [p 139]. This leads us to believe that in order to make a system fully compatible to the user, the user's culture should be accommodated into the design of the interface. However, not everyone in the HCI community seems to agree: The literature contains as many arguments against as supporting the accommodation of culture into user interface design. Those that are in support of culture do not seem to agree on how to implement such accommodation. Those opposing views highlight that the problem with cultural issues in the HCI field is that 'the theoretical foundation for cultural influences on interface design is confused and not yet at the level which would allow designers to predict where culture will have an influence on products, or how these differences can be addressed' [Dunckley and Smith, 2000, p 3].

To better understand how culture influences HCI, we begin this section by reviewing different definitions, metamodels and models of culture. We then analyse the arguments in favour of and against the relative importance of culture in HCI. Finally we review the approaches to culturalization and present a metamodel for accommodating both subjective as well as objective cultural variables into the design of interfaces.

2.5.1 Definitions of Culture

There are many definitions, approaches and models in the literature, but there is no agreement on a specific definition of culture [Hoft, 1996; Ciborowski, 1979; Vohringer-Kuhnt, 2003]. Some examples of such definitions include:

- 'Culture can be viewed as an organized collection of rules focusing on the ways in which members of a population should communicate, think, and interact with one another and their environments' [LeVine, in Ciborowski, 1979, p 102]
- Culture is defined as 'behaviour typical of a group or class of people' [Webster Hypertext Interface, 1995]
- Culture is conceptualized as a 'system of meaning that underlies routine and behaviour in everyday working life' [Bodker and Pederson, 1991, p 122]
- Culture 'includes race and ethnicity as well as other variables and is manifested in customary behaviours, assumptions and values, patterns of thinking and communication style' [Borgman, 1986, p 49]
- 'Culture is communication, and communication is culture' [Hall, 1959, p 186].
- Culture is 'the collective programming of the mind that distinguishes the members of one group or category of people from another, where the mind stands for thinking, feeling and acting, with consequences for beliefs, attitudes and skills' Hofstede, 2001, p 5]

Instead of trying to produce a general definition, it is more important to find a definition of culture that serves this specific area of research [Honold, 2000]. Many of the above definitions refer to culture as influencing the way in which communication takes place. As noted earlier, using the computer to perform tasks requires communication between the user and the system, particularly when using an interactive system. Consequently, for the purposes of this research, we define culture as the patterns of thinking, feeling and acting that influence the way in which people communicate amongst themselves and with computers.

2.5.2 Metamodels of Culture

Culture is a complex concept that cannot just be described: it has to be interpreted [van Peursson, in Evers and Day, 1997]. This implies that there are various dimensions of culture that should be considered. Metamodels of culture provide a high-level view of the overriding philosophies surrounding the concept of culture, by defining different layers of culture [Hoft, 1996]. Hoft [1996] provides a useful overview of four of the metamodels that have been identified in the literature, namely (1) the Iceberg Model, (2) the Onion Model, (3) the Pyramid

Model and (4) the Objective and Subjective Culture Model. This overview is summarized and paraphrased here.

2.5.2.1 The Iceberg Model

The Iceberg model is a popular model in cross-cultural training, but its origins are difficult to trace. It identifies three layers of culture: surface, unspoken rules and unconscious rules.

- The surface layer comprises cultural characteristics that are visible and obvious. They include number, currency, date and time formats; language, font direction and character sets
- Unspoken rules relate to a specific context or situation such as business etiquette and protocol. The cultural characteristics relevant to this layer are somewhat obscured as the context needs to be identified before the unspoken rules can be understood.
- Unconscious rules are out of conscious awareness and are therefore difficult to study. Examples include nonverbal communication, a sense of time and physical distance, and the rate and intensity of speech.

2.5.2.2 The Onion Model

The Onion model was developed by Trompenaars [1993] and identifies three layers of culture: the outer layer, the middle layer and the core.

- The outer layer consists of explicit products and artifacts of culture. Trompenaars suggests that the outer layer consists of of the first things we encounter at a cultural level, and includes aspects such as language, food, buildings, monuments, agriculture, fashions and art.
- The middle layer defines norms and values. Norms are principles of right and wrong that are shared by a group of people. Values define what is good and bad.
- The core of culture is implicit and consists of how people adapt to their environments. The core consists of the basic assumptions of human existence.

2.5.2.3 The Pyramid Model

The Pyramid Model was introduced by Hofstede [2001] and identifies three layers of culture: personality, culture and human nature. These three layers focus on understanding the origin of culture and why it is unique in human mental programming, and include.

- Personality: is specific to an individual and is both learned and inherited.
- Culture: is specific to a group or category of people. In contrast to personality, it is learned, not inherited.
- Human nature: is what is common to all human beings. Unlike culture and personality, human nature is inherited, not learned. It is passed down through the generations via DNA.

2.5.2.4 The Objective Culture and Subjective Culture Model

The Objective Culture and Subjective Culture Model was developed by Stuart and Bennett [in Hoft, 1996] and identifies only two layers of culture, objective culture and subjective culture. Objective culture is the 'institutions and artifacts of a culture, such as its economic system, social customs, political structures and processes, arts, crafts and literature' [p 43]. These characteristics have been further classified into levels of culture [Hofstede, 2001], and include:

- a national level according to one's country;
- a regional and/or ethnic and/or religious and/or linguistic affiliation level, as most nations are composed of culturally different regions and/or ethnic and/or religious and/or language groups;
- a gender level, according to whether a person was born as a girl or as a boy;
- a generation level, which separates grandparents from parents from children;
- a social class level, associated with educational opportunities and with a person's occupation or profession; and
- for those who are employed, an organizational or corporate level according to the way employees have been socialized by their work organization.

Objective culture is visible, easy to examine and tangible, as it is represented in terms of text orientation, date and number formats, colour and language [Hoft, 1996]. Comparing this layer to the layers identified in the other models, there is a clear correlation to the outer layer of the Onion model and the surface layer of the Iceberg model.

In contrast, subjective culture is 'the psychological features of a culture, including assumptions, values, and patterns of thinking' [Hoft, 1996, p 43]. Subjective culture is difficult to examine because it operates outside of conscious awareness, for example, in the way in which people accept or reject uncertainty [Hofstede, 2001], similarities and differences in power and authority [Victor, 1992; Hofstede, 2001], and the amount of emotions that people express when dealing with others [Trompenaars, 1993]. Subjective culture is therefore evidently related to the middle and core layers of the Onion model, the unspoken rules and unconscious rules layers of the Iceberg Model, and all the layers in the Pyramid model.

For the purposes of this research, culture is categorized into subjective culture and objective culture. Objective culture is abstract, because it is an externalization of subjective culture. Subjective culture is what is real and concrete. However, objective culture tends to be treated as more real and concrete than its source, subjective culture [Stewart and Bennett, in Hoft, 1996].

2.5.3 Models of Culture

The metamodels of culture described above form the basis for the development of different models of culture. These models provide a more detailed view of culture, by identifying a number of cultural dimensions that are used to organise cultural data [Evers, 2001]. Hoft [1996] identified four models of culture, developed by Victor [1992], Hall [1959], Trompenaars [1993] and Hofstede [2001], which are summarized in Table 2.2.

Victor [1992]	Hall [1959]	Trompenaars [1993]	Hofstede [2001]
• Language • Environment and Technology • Social Organisation • Contexting • Authority Conception • Nonverbal Behaviour • Temporal Conception	• Speed of Messages • Context • Space • Time • Information Flow • Action Chains	• Universalism vs. Particularism • Neutral or emotional • Individualism vs. Collectivism • Specific vs. Diffuse • Achievement vs. Ascription • Time • Environment	• Power Distance • Masculinity vs. Femininity • Individualism vs. Collectivism • Uncertainty Avoidance • Time Orientation

Table 2.2: Examples of Cultural Models and Their Dimensions

Each of these cultural models is now discussed in terms of their primary focus and their dimensions.

2.5.3.1 Victor's Model of Culture

Victor's [1992] model focuses on determining the aspects of culture that are most likely to affect business communications.

- Language: this dimension refers to objective culture and includes the degrees of fluency, accents and regional dialects and how they affect business communications. The pros, cons and misperceptions of using English as the international language of business are also discussed.
- Environment and technology: refers to the larger issues of how geography, population, physical space and perceptions of technology influence business communication.
- Social organisation: includes the effects of educational, economic, social, political and religious systems on business communication. These issues are predominantly related to objective culture.
- Contexting: this is the same dimension that appears in Hall's cultural model and is discussed in more detail in section 2.5.3.2.
- Authority conception: this dimension considers differences and similarities in power, authority and leadership, and is similar to the dimension of time orientation discussed under Hofstede's [2001] cultural model in section 2.5.3.4.
- Non-verbal behaviour: this is a broad category for many types of non-verbal behaviour, including active behaviour such as movement, sound, eye and touching, and passive

behaviour such as the use of colours, symbols and smells.

- Temporal conception: this is similar to Hall's [1959] dimension of time and includes polychronic and monochronic time, particularly in terms of scheduling.

Language, environment and technology and social organization all relate to objective cultural aspects, while contexting, nonverbal behaviour, authority conception and temporal conception are considered subjective culture. Victor's model thus includes both subjective and objective cultural dimensions.

2.5.3.2 Hall's Model of Culture

Hall [1959] describes culture as a selective screen through which we see the world, and believed that basic differences in the way that members of different cultures perceived reality were responsible for the miscommunications of the most fundamental kind [Smith et al., 2004]. Consequently, Hall's model focused on determining what releases the right response rather than what sends the right message. Hall distinguished cultures on the basis of a way of communicating along a dimension from 'high-context' to 'low-context'.

- Context: context refers to the amount of information that is in a given communication as a function of the context in which it occurs. A high context communication is one in which little has to be said or written because most of the information is either in the physical environment or within the person, while very little is in the coded, explicit part of the message [Smith et al., 2004]. People who know each for a long time generally tend to use high context communication. In contrast, low context communication is similar to interacting with a computer: if the information is not explicitly stated, the meaning is distorted.
- Speed of messages: also known as the speed velocity continuum, this refers to the speed with which people decode and act on messages. Some cultures prefer slow messages, such as deep relationships, books, poetry and television documentaries, whilst other cultures prefer fast messages, such as headlines, cartoons, television commercials and newspaper headlines
- Space: Hall identified four different senses of space, or invisible boundaries, namely territoriality, personal space, multisensory space and unconscious reactions to spatial differences. Territoriality includes 'ownership' and communicates power and authority. For example, in the United States, those with power and authority typically have the largest and most lavish offices that are often located on the highest floors of a building. Personal space is where different cultures have unspoken rules about how close one is allowed to get to another person. Multisensory space refers to the invisible boundaries placed on the five senses, and relates to the unconscious rules about what is too loud and intrusive. The boundaries of multisensory space are linked to context, as high context cultures expect loud conversations to occur. In contrast, low context cultures consider loud conversations to be an infringement on another person's personal space.

Unconscious reactions to spatial differences refer to the response that is invoked as a result of the distance that a person keeps while having a conversation. This is related to personal sense of space, for example, if a person is having a conversation at a distance closer or further away than expected, an unconscious and negative message is being sent to that person.

- Time: Hall identifies time in terms of polychronic and monochronic time. Polychronic time is where people tend to do many things at once, and generally portray characteristics of being high-context, committed to people and relationships, change plans often and are easily distracted. In contrast, monochronic time is where people do one thing at a time, and portray characteristics of low context, commitment to the job, and adhere strictly to plans. Hall also talks about past- and future-oriented cultures which is similar to Hofstede's time-orientation dimension and will be discussed in more detail in section 2.5.3.4.
- Information flow: refers to how long it takes for a message to travel to, and release the required response from, the intended recipient. Once again this dimension is related to context, as information flow in high context cultures tend to be fast and free. In contrast, information flow in low context cultures tends to be hampered by bureaucracy and rigid procedures.
- Action chains: action chains are a sequence of events that lead to the accomplishment of a goal, and refer to the important rules governing the way in which tasks are executed.

Information flow and action chains are seen as objective culture, whilst the other dimensions that comprise this model of culture relate to subjective culture. Like Victor's model, Hall's model thus incorporates both subjective as well as objective cultural dimensions.

2.5.3.3 Trompenaars' Model of Culture

Trompenaars' [1993] focus for his model was to determine the way in which a group of people solves problems [Smith et al., 2004]. He identifies seven dimensions of culture that characterize the way cultures solve problems. These are described below [Hoft, 1996; Gould et al., 2000].

- Universalism vs. particularism: universalists are rules-based, therefore, in a serious situation involving another person, universalists tend to apply rules of morality, ethics and what is good and right regardless of their relationship with the other person. In contrast, particularists base their solution to the problem on the relationship that they have with the other person, and are prepared to break the rules if necessary.
- Neutral or emotional: this is a measure of the range of emotions that people express when dealing with others in a business environment. Neutral people keep their emotions in check and focus on argument. In contrast, emotional people show their reactions and expect emotional reactions in return.
- Individualism vs. collectivism: refers to the balance that is perceived to be appropriate

between the needs of individuals and groups [Gould et al., 2000]. This dimension also appears in Hofstede's model and is discussed in more detail in section 2.5.3.4.

- Specific vs. diffuse: a specific value orientation is one where public and private life and personal space is compartmentalized. It measures the range of involvement that people have with others in their lives. In specific-oriented cultures, there is a clear division between business relationships and personal friendships. In diffuse cultures, business communications require the development of strong personal relationships such as liking and trust, before co-operation can begin.
- Achievement vs. ascription: refers to how status is accorded to others [Gould et al., 2000]. Achievement oriented cultures achieve status by 'doing', in other words, through individual achievements. In contrast, ascribed status comes from 'being' in a certain role, and is often based on gender, age, social connections and education. .
- Time: this dimension is present in all four of the cultural models discussed here. It includes Hall's definition of polychromic and monochromic time, as well as a culture's attitude towards the past, the present and the future and the relationship of the three to each other.
- Environment: in contrast to the similar dimension found in Victor's [1992] model, the environment dimension in Trompenaar's model relates to the attitudes that different cultures have towards their ability to control nature. The main difference between cultures is in the belief that it is either worthwhile to try to control nature (such as Brazil, Portugal and China), or not worthwhile (such as Japan, Switzerland and Singapore).

In contrast to the models discussed previously, Trompenaars' model incorporates only subjective culture as all of the above variables relate to beliefs, and values, rather than artifacts and institutions.

2.5.3.4 Hofstede's Model of Culture

In keeping with his definition of culture described earlier, Hofstede [2001] focuses his model on determining the patterns of thinking, feeling and acting that form a culture's mental programming. He conceptualized culture as 'programming of the mind' in the sense that certain reactions were more likely in certain cultures than in others, based on differences between the basic values of the members of different cultures [Smith et al., 2004].

In 1967, Hofstede conducted an initial survey of IBM's middle management in 70 countries in an attempt to identify the cultural dimensions that can be used to distinguish among different cultures [ITIM, 2003]. Based on the results obtained from 40 of the countries surveyed, he identified four cultural dimensions, namely power distance, uncertainty avoidance, masculinity/femininity, and individualism/collectivism. The scope of the survey was later extended in a second study to include 50 countries and three regions, the results of which validated the results of the earlier study. In the late 1980's, Hofstede identified a fifth dimension based on Chinese values, namely time orientation. Subsequent studies on commercial airline pilots and students in 23 countries, up-market consumers in 15 countries, elites in 19 countries and civil service managers in 14 countries further validated the results of Hofstede' s initial surveys [ITIM, 2003].

All five of Hofstede' s dimensionsrelate to subjective culture. Each of these dimensions is a dichotomy, in that there are two opposing sides to each dimension. Many of the dimensions in his model also appear in the other three models discussed above. An overview of the cultural dimensions is presented here [Evers, 2001; Hofstede, 2001; Hoft, 1996; Marcus, 2000].

- Power distance: refers to the extent to which less powerful members expect and accept unequal power distribution within a culture. The two opposing sides to this dimension are high power distance and low power distance.
- Uncertainty avoidance: relates to the way in which people cope with uncertainty and risk. The two opposing sides to this dimension are high uncertainty avoidance and low uncertainty avoidance.
- Masculinity vs. femininity: refers to gender roles, as opposed to physical characteristics, and is commonly characterized by the levels of assertiveness or tenderness in the user. The two opposing sides to this dimension are masculinity and femininity.
- Individualism vs. collectivism: refers to the role of the individual and the group, and is characterized by the level of ties between an individual in a society. The two opposing sides to this dimension are individualism and collectivism.
- Time orientation: refers to people's concerns with the past, present and future. The two opposing sides to this dimension are short-term orientation and long-term orientation.

A more detailed discussion of Hofstede's cultural dimensions is presented in Chapter 4, as this model was used as the theoretical foundation for an experiment that explored the relationship between subjective culture and usability.

2.5.4 Influence of Culture in Human-Computer Interaction

We argue that as culture influences the way in which people interact in general, culture will also influence the way in which people will interact with computers. Using interactive systems to perform tasks requires communication between the system and the user. People learn patterns of thinking, acting and communicating from living in a specific social environment, normally typified by national culture [Massey et al., 2001]. As such, culture partially pre-determines a person's communication preferences and behaviours. Communication style, which reflects how a person sends and interprets messages, represents the overall patterns and values of a culture. As the user interface is the means by which the user and the computer interact [Dix et al., 1998], it stands to reason that the interface should facilitate users to use their particular communication styles [Massey et al., 2001]. Consequently, global interfaces need to accommodate a diversity of communication styles to provide support for the cultural diversity of the users.

Diversity in culture is particularly relevant where global interfaces take the form of websites [Jagne et al., 2004; Barber and Badre, 1998; Chau et al., 2002; Del Galdo and Nielsen, 1996; Yeo, 1996; Marcus, 2001; Smith and Chang, 2003; Hall, et al., 2003]. The advent of the World Wide Web (WWW) has resulted in a fundamental technical context of use that now needs to be taken into consideration, namely the difference between traditional software applications and web applications. As at the end of September 2003, there were over 43 million websites connected to the Internet [Lee and Kim, 2004], with a significant number of these sites being designed for organizations that use the WWW as a medium for international communications and transactions [Chau et al., 2002].

As with traditional user interfaces, the web interface should facilitate ease of use and learning of the system, ensuring that the information is conveyed in a cognitively efficient way [Chau et al., 2002]. However, there are additional challenges to the design of usable web interfaces that are not present in traditional user interfaces. Nielsen [1997] suggests that these differences relate primarily to the designer having to 'give up full control and share responsibility for the user interface with the users and their client hardware/software'. He identifies three main differences relating to control and shared responsibility, namely device diversity, navigation control and part of a whole:

(1) Device diversity: relates to having to set aside all assumptions about the system when the designer is designing a web interface. In traditional GUI interfaces, the designer knows the system, the fonts, the screen size and the bandwidths that the user will be using. These can differ significantly when users are accessing a web interface, as a result of the diversity of technological devices that are available.

(2) Navigation control: relates to the level of control that can be designed into the interface. With a GUI interface, the designer can control where user can go, using, for example, disabled menus and modal dialog boxes. With a Web interface, the navigation is user controlled, in that the user can go where designer never intended. For example, the user can access a page of the website through a search engine. This also results in users wondering if they have missed out on information that may be available on other pages in the site.

(3) Part of a whole: relates *to how* users perceive different applications to be different, and therefore expect differences in features, layout and conceptual models. Although each website is actually a different application, users do not often perceive them to be different, and are not expecting differences and unfamiliarity between sites.

From a cross-cultural usability perspective, the primary difference between traditional software and web-based interfaces is that websites are constantly addressing different cultural audiences simultaneously [Chau et al., 2002]. Within the global information technology (IT) environment, cross-cultural usability of websites is about making websites an effective means of communication between a global web site owner and a local user [Smith et al., 2004]. Because users differ across regional, linguistic and country boundaries, their expectations of websites are driven primarily by their local cultural perspectives. Consequently, user reactions become more predictable and understandable when the user's cultural perspective is taken into account [Barber and Badre, 1998; Massey et al., 2001]. Websites need to display 'culturability', that is, designing the interface to accommodate the cultural preferences and biases to increase the usability of the interface and the product [Barber and Badre, 1998].

However, there are still many arguments against the influence and consequent importance of culture in HCI. The first argument is based on grounds of cultural convergence [Bryan et al., 1994; Norton, 2001; Light in UsabilityNews, 2003]. Cultural convergence theories suggest that cultures have a tendency to become similar due to technology and global industrialization influences [Bryan et al., 1994]. 'Culture may be embedded in technology, but technologies change culture' [Light in UsabilityNews, 2003]. Computer users are increasingly exposed to influences from the global economy, resulting in the transfer of knowledge, behaviours and standards between cultures. This suggests that all computer users, irrespective of their country of origin, belong to a 'computer / technology sub-culture' that uses a specific language, set of symbols, values and protocols for behaviour' [Norton, 2001]. The existence of such a sub-culture essentially refutes the need for accommodating both objective, as well as subjective, cultural differences. For example, Mrazek and Baldaccini [1997] points out that the use of a trashcan to signify the deletion of files for European and Asian users is often thought to be a mistake, as the users may not recognize the icon. However, if users have previously been exposed to this icon, they will not only recognize it, but will also expect it. Therefore, exposure to another culture may cause expectations, which if not met, will confuse the users more than attempts to accommodate their cultural diversity.

The convergence argument is counteracted by the theory of cultural divergence. This theory points out that conformance to global pressures will not necessarily override specific cultural values and practices [Bryan et al., 1994]. Instead, continuous pressures from other cultures will result in 'an outburst of attitudes of defense of national and regional identities, and manifestations of the fear of a mixing of races, religions, customs and habits' [De Souza and Dejean, 1999]. Consequently, sub-cultures are important, but they are superceded by the culture of the country [Marti and Muller, 2003]. This is supported by Child's [in Bryan et al., 1994] hypothesis that convergence and divergence should be synthesized: cultural convergence is most likely to be found in work structures, whereas divergence is more apparent in work practices. For example, Bryan et al.'s [1994] study found that the work done by information systems professionals is conceptually similar across cultures (convergence), but that technology does not always eliminate differences in cultural beliefs and practices (divergence).

A second argument against the importance of culture in HCI relates to the purpose for which the interface is being used. Chau et al. [2002] identified four main purposes for web interfaces: information searching, hobbies, social communication and e-commerce. According to the Wharton Virtual Test Market study [Chau et al., 2002; Fitzgerald, 2004], the most important factor affecting on-line buying behaviour is how much individuals used the web to search for product information. Therefore, Fitzgerald [2004] contends that the factor that most influences attitudes and behavioural responses to websites is the purpose for which the site is being used, rather than culture. However, Chau et al. [2002] point out that it is the culture of the users that dictate the different purposes for which websites are used. Their study found that US users used websites primarily for searching for information, whereas Hong Kong users viewed the Internet as a social interaction mechanism. Furthermore, the different purposes may lead users to have different impressions of the same sites. For example, website interfaces for US users that do not provide an efficient mechanism for searching for product information will not be perceived as very useful. In contrast, Hong Kong users would be more accepting of website interfaces that provide a virtual community-like environment. Ultimately then, it does not matter what the purpose of the site is: 'disregard cultural differences and your site will be doomed' [Forbes, 2001, in Chau et al., 2002]. Consequently we agree that cultural differences do matter [Massey et al., 2001], and that it is necessary to accommodate for cultural diversity in software products and Internet sites to meet the needs of the culturally diverse market.

2.5.5 Approaches to Culturalisation

Culturalisation, or preparing a product for use by diverse cultures, requires two steps: internationalization and localization. Internationalisation involves identifying the culturally specific elements of the product, and localization involves substituting those culturally specific elements with a local content [Russo and Boor, 1993].

Traditionally, the approaches to culturalisation seemed to focus primarily on objective cultural issues rather than subjective culture. The objective cultural approach concludes that, when dealing with human-computer interaction, meaning is the central issue in culture [Bourgess-Waldegg and Scrivener, 1998]. Supporters of this approach suggest that designers need only cater for cultural diversity by ensuring that the intended meaning of user interface representations, such as symbols, icons and language, are translated to suit the target cultures, so that they are understood correctly. Thus, this approach is based on the premise that it is the objective, rather than the subjective, cultural aspects that are important. The culturalisation process has concentrated primarily on the translation of the objective cultural aspects [Dunckley and Smith, 2000], such as language and date and time formats [Bourges-Waldegg and Scrivener, 1998; Russo and Boor, 1993], to avoid the types of amusing, but potentially harmful misunderstandings listed in Table 2.3.

Name of Product	Intended Meaning	Interpreted Meaning
English name of a USA product	Clairol Mist Stick	Piece of Manure (German)
English name of a USA product	Matador (AMC auto)	Killer Auto (Spanish)
English name	Guess Jeans	Vulgar / Low-class / Ugly Jeans (Japanese)
Product Term	Credit Card	Guilt Card (German)

Table 2.3: Misunderstandings Caused by Language Translations [Hair, et al., 2003]

However, researchers are finding this method inadequate [Jagne et al., 2004]. It has been argued that whilst objective culture is important, it is also necessary for the interfaces to reflect the values, ethics and morals of the target users [Russo and Boor, 1993], in order to make the users more comfortable and accepting of the interfaces. These aspects relate to subjective culture [Dunckley and Smith, 2000], and go beyond the 'surface manifestations of culture that have been widely accepted' [Kersten et al., 2000, in Smith et al., 2004, p 89]. Carey [1998] points out that there is a difference between comprehension and acceptance. She suggests that it is not enough just to translate the representations, but that it is equally important that *other cultural conventions* feel comfortable and recognizable to the user, thereby increasing the usability of the interface. Russo and Boor [1993] expand on this suggestion, pointing out that in addition to language, interfaces need to reflect the values, ethics and morals of the target users [Russo and Boor, 1993].

More specifically, there is an inseparable link between knowledge and culture [Ciborowski, 1979]. For example, what may be believed to be an irrefutable fact in one culture may be deemed to be groundless superstition in another culture. Members of a culture may refuse to view the particular beliefs and/or traditions of a different culture as worthy of being called knowledge. Therefore, any investigation that attempts to assess the knowledge of a group, without understanding and assessing the culture of that group, is inherently confounded [Ciborowski, 1979]. In the context of user interface design, knowledge is a determinant of user performance. As culture influences knowledge, and knowledge influences performance, this implies that culture also influences usability and performance. Del Galdo and Nielsen [1996] clearly support this by pointing out that there are three levels of internationalization, namely:

(1) Displaying the native language, character set and notations.
(2) Translating the user interface and documentation so that it is understandable and usable.
(3) Matching the user's cultural characteristics, which goes beyond avoiding offensive icons and must accommodate the way business is conducted and the way people communicate.

Essentially, this approach is based on the premise that culture is about how individuals behave and respond, their beliefs and values, and therefore it is also necessary to reflect subjective culture in the design of interfaces [Dunckley and Smith, 2000]. Smith et al. [2004] also support this view for the culturalisation of website interfaces: they believe that 'the process of internationalization and localization for web-based systems involve an analysis of both objective and subjective cultural issues' [p 83].

This highlights that culturalisation should be addressed from both sides of the cultural spectrum: subjective cultural issues should be accommodated using a top-down approach, whilst objective cultural issues should be dealt with using a bottom-up approach. We suggest that subjective culture provides the context or conceptual model within which the interface is designed, thus forming the higher-level, semantic layer of the interface. The more detailed syntactic level should be governed by objective culture, covering issues such as language, text and date formats and content. This is depicted in the metamodel illustrated in Figure 2.2

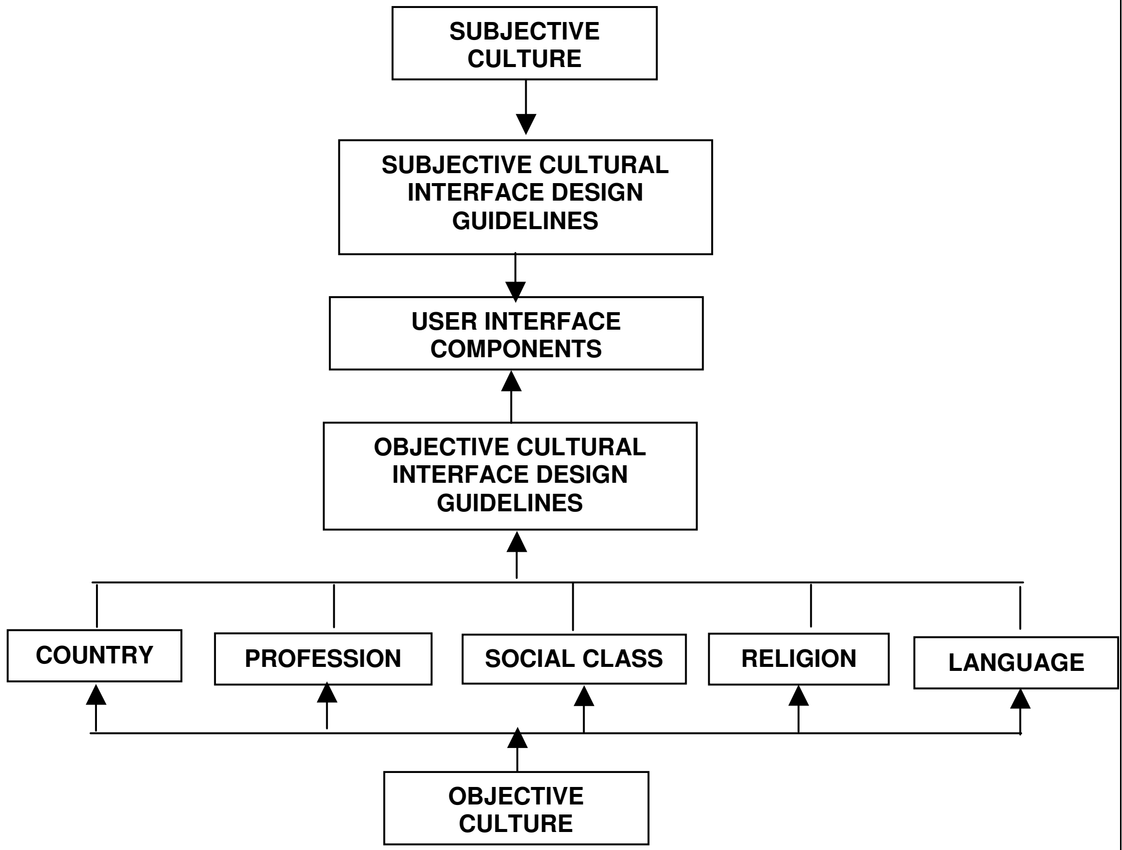

Figure 2:2: Proposed Metamodel for Accommodating Culture into User Interface Design

2.6 Subjective Cultural Interface Design Guidelines

The metamodel proposed in Figure 2.2 highlights the need for subjective cultural interface design guidelines in order to accommodate subjective culture into the design of the interfaces. According to Fitzgerald [2004], these guidelines are based on one of four models currently being used for managing the subjective aspects of cross-cultural website design: namely (1) cultural behaviour models, (2) models based on Activity Theory, (3) cultural markers model and (4) cultural dimension models.

The cultural behaviour model is based on a study of different on-line behaviours of website users from the US and Hong Kong [Chau, et al., 2002]. It was reported that different cultures use the Internet for different purposes, such as social communication, e-commerce, hobbies and information search. They suggest that culturability stems from identifying the purpose for which a particular culture uses the Internet, and then ensuring that the site concentrates on providing features to assist the achievement of this purpose. Fitzgerald [2004] points out that this approach is relatively young, and further research is needed to determine how useful it turn out to be.

Activity Theory is a socio-cultural theory of human activity [Fitzgerald, 2004]. The basic model describes subjects attempting to achieve goals (objects) via activities. Honold [2000] contends that this theory offers a useful approach for integrating cultural influences in HCI, as the theory defines not only the interaction between humans, tools and the environment, but also the interaction between individual activity and sociocultural context. However, Fitzgerald [2004] points out that the use of activity theory in HCI is relatively new, and further, that its roots in Marxist theory may limit its use for creating business websites.

The third set of guidelines is the cultural markers model [Sun, 2001; Barber and Badre, 1998]. Cultural markers are interface design elements and features that are prevalent, and possibly preferred, within a particular cultural group [Barber and Badre, 2003]. Also referred to as 'cultural attractors', cultural markers are interface design elements that reflect the signs and their meanings that match the expectations of the local culture [Smith et al., 2004]. Examples include national symbols such as flags, colour, colour combinations and leaders, as well as trust signs, language cues, currency formats and navigation controls, that together create a 'look and feel' to match the cultural expectations of users for a particular domain. Cultural markers aim to 'address quite specific cultural aspects directly, rather than attempting to build upon generic cultural differences and models' [Smith et al., 2004, p 73]. Therefore, the cultural markers model offers a 'bottom-up' approach to culturalisation which complements the 'top-down' approach of cultural models. Fitzgerald [2004] believes that the cultural markers model shows the most promise, even though it is relatively young and initial studies are not very supportive.

The fourth set of guidelines is based on cultural dimension models, such as those discussed in section 2.5.3. Interest in cultural dimension models arose as a result of a need for models of cross-cultural software development [Fitzgerald, 2004]. However, the use of cultural models has been criticized based on discomfort with stereotypes [Light, 2003; Bourges-Waldegg and Scrivener; 1998] and the inconclusive results obtained from studies done to evaluate and apply these models to HCI [Fitzgerald, 2004]. A stereotype is defined as a 'fixed notion about persons in a certain category, with no distinctions made among individuals' [Hofstede, 2001, p 14]. An example of inconclusive results is the study conducted by Evers and Day [1997] that found that 'Asians prefer to instruct computers using detailed commands, see uncertainty as threatening, and prefer text-based rather than symbolic interfaces', which is not what one would expect from their subjective cultural profile. Bourges-Waldegg and Scrivener [1998] consider the use of cultural models inadequate because they are generalizations and therefore do not take into account the context in which they are applied. In addition, not everyone fits into the 'cultural mean', because there is a wide range of individual differences [Light. 2003; Jagne et al., 2004].

In particular, Hofstede's cultural model, which is the most often quoted of the four cultural models discussed previously [Smith et al., 2004], has endured much criticism from the HCI and usability-related community. For example, Hall [2001, in Jagne et al., 2004] says that 'it would be wrong to take the characteristics of Hofstede and attempt to deduce from these how a particular cultural group would respond to a particular technology'. Hofstede's model has been considered too stereotypical [Bourges-Waldegg and Scrivener, 1998] or rigid [Nocera and Hall, in Jagne et al., 2004]. In addition, previous attempts to apply Hofstede's model to usability has also resulted in conflicting and therefore inconclusive findings. For example, Gould et al. [2000] found that Malaysian websites contain links on the home page to website administration, which correlates well with the high power distance reported [Hofstede, 2001] for Malaysia. However, this does not explain why low power distance countries such as the US also contain such links on their websites. In contrast, the Forer and Ford [2003] study (discussed in more detail below) found that accommodating for the user's cultural profile enhanced performance. Consequently, until better proof of their relevance to website design is provided, Fitzgerald [2004] suggests that cultural dimension models should be used with care.

In contrast to the arguments put forward against the use of cultural models in general and against the use of Hofstede's [2001] cultural model in particular, we believe that the literature provides sufficient theoretical evidence to support the use of Hofstede's cultural model as the basis for managing the subjective cultural aspects of interface design. Analysis of this model indicates that accommodating these dimensions into the design would decrease the cognitive load, thus enhancing usability and performance. In addition, a number of studies in the literature suggest that these dimensions are related to usability. This theoretical evidence is presented in the section that follows.

2.6.1 Influence of Hofstede's Cultural Dimensions on Cognitive Load

Users interpret new information on the basis of their existing mental models [Honold, 2000]. Hofstede's cultural model is focused on determining the patterns of thinking, feeling and acting that form a culture's mental model [Hoft, 1996]. This leads us to believe that Hofstede's cultural dimensions do influence the cognitive resources required for performance to take place. In terms of the attraction and identification processes, we suggest that the metaphors used should be pertinent to the cultural dimension of the user, particularly if the metaphor is used as a way of representing a specific function of the interface. For example, high power distant users would be more attracted to metaphors that focus on expertise, authority, official stamps and logos, whereas low power distant users would be more attracted to metaphors that focus on images of equality, such as playgrounds and public places [Marcus, 2002].

Similarly, we can see that the cultural dimensions influence the reasoning, problem solving and decision-making processes by:

- Influencing the user's perception of concrete data: as we explained previously, concrete data may alleviate the complexity of the problem, thus facilitating problem solving during the analysis phase of the cognitive process. We suggest that the user's perception of concrete data can be influenced by the user's cultural dimension. For example, a high uncertainty avoidant user prefers detailed explanations and information, whereas a low uncertainty avoidant user prefers summarised data [Marcus, 2002].
- Maintaining the user's concentration during reasoning and problem solving: reasoning and problem solving require concentration. Any distractions could cause a loss in concentration, resulting in a need to start again, thus reducing speed, and sometimes accuracy, of performance. Slow response times also cause the user to become distracted from their task through boredom or irritation, causing loss of concentration. This is exacerbated in users that are naturally impatient or wanting to complete tasks quickly, such as masculine and short-term oriented users [Marcus, 2002].
- Reducing anxiety: anxiety reduces the size of memory, as the user is partially absorbed by concerns that are beyond the problem-solving task [Shneiderman, 1998]. For example, high uncertainty avoidant users are generally more emotional and stressed than low uncertainty avoidant users, and may be so anxious about having to learn how to navigate through the interface that they are not concentrating on the problem itself. As another example, high power distant users may be worried about disappointing their superiors, thus putting themselves under more pressure and causing additional anxiety [Marcus, 2002].

Hofstede's cultural dimensions also appear to affect the way in which users will respond to messages. For example, collectivist users would be most uncomfortable with having to express personal opinions, while high power distant users would not wish to express an opinion that is in direct contradiction to their superiors. Similarly, individualist users would not be comfortable with having to provide personal information, but would find that being unable to express a personal opinion unacceptable [Marcus, 2002]. This, in turn, will affect the functionality provided in the interface. For example, a study conducted on students' use of a hypertext product to analyze a poem found that students should be able to add their own viewpoints to the system in some cultures (low power distant and individualist), but in others (high power distant and collectivist) this would be unacceptable [Russo and Boor, 1993]. Because the product did not provide a function that allowed students to add their viewpoints, the system was considered unacceptable to some cultures. Therefore, the functionality provided in the interface should be consistent with the users' cultural dimensions, thus adhering to the related principle of task conformance, and providing users with an appropriate mechanism with which to respond. This will reduce the possibility of a loss of speed and accuracy, increased training time, and reduction in user satisfaction.

2.6.2 Influence of Hofstede's Cultural Model on Usability

Further evidence that Hofstede's model is relevant to cultural diversity in HCI was found in the literature. A limited number of studies focusing on the relationship between Hofstede's cultural model and usability were found. An analysis of these studies suggests that Hofstede's model is related to usability from three different perspectives, namely user acceptance, context of use and objective usability. These perspectives are discussed next.

2.6.2.1 The User Acceptance Perspective

User acceptance is a distinct area of HCI research that explores the reasons why people accept or reject information technology. Many models have been developed for this purpose, including the Technology Acceptance Model [Davis, 1989]. Many of the studies relating to Hofstede's model of culture identified in the literature focus on examining the role of subjective culture on user acceptance of information technology, rather than on usability. These studies are based on the Technology Acceptance Model (TAM), which is depicted in Figure 2.3.

The components of TAM are defined as follows [Davis, 1989]:

- Perceived ease of use is defined as the degree to which a person believes that using a particular system would be free from effort. Ease of use was seen to include physical effort, mental effort, and ease of learning.
- Perceived usefulness is defined as the degree to which a person believes that using a particular system would enhance his or her job performance. Usefulness was found to incorporate the categories of job effectiveness, productivity, time saving and importance of the system to the user's job.

- Behavioural intention is the measure of the strength of one's intention to perform a specified behaviour.
- Attitude is defined as the individual user's positive or negative feelings about performing the target behaviour
- Actual use is measured in terms of the user's frequency of system use (how often) and the volume of system use (how much).

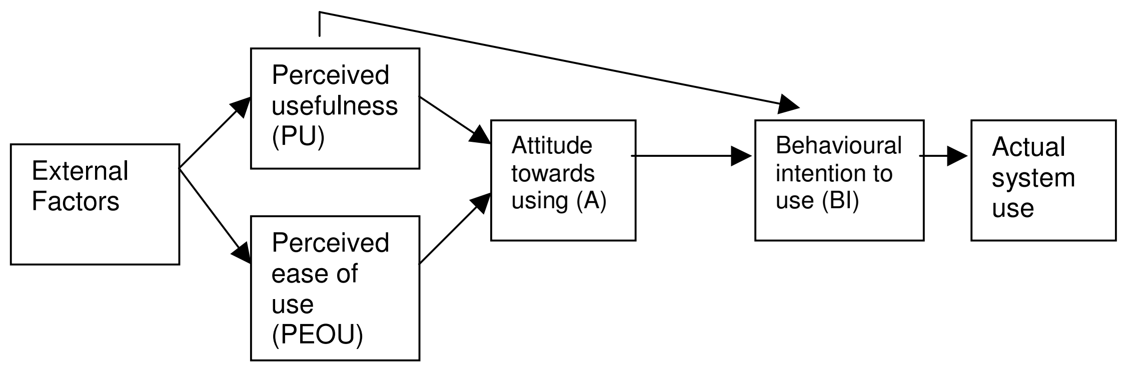

Figure 2.3: The Technology Acceptance Model [Davis, 1989]

As illustrated in Figure 2.3, this model identifies perceived ease of use and perceived usefulness as the two basic determinants of user acceptance. Acceptance, as defined by the TAM literature, is therefore based on user perceptions of usability, rather than actual performance. The question that arises is whether 'acceptance' is related to usability, or whether these two concepts should be treated as two mutually exclusive attributes of a product.

One school of thought contends that acceptance is not related to usability as defined in the usability engineering literature, because usability tends to be a performance-based concept [Bietz and Olson, 2001]. The primary concerns include the number of errors made by users, the amount of time that users take to complete a task, or the percentage of users who are able to successfully complete tasks. In contrast, TAM was developed to predict user opinions about the product before the product was actually purchased and used. In most of the research on this model, the focus has been on understanding the links between perceptions and intentions ('attitude towards using') and between intentions and actual use. A typical experiment used in the TAM research involved demonstrating new software for a user, and then asking the user to make value judgments about whether or not they are going to use the product, based on their perceptions of the usefulness and ease of use of the product. This approach does not allow for usability in terms of accuracy and speed to be measured because the user has not yet used the product. In addition, it has been argued that more meaningful perceptions of usability are developed only after the user has had hands-on experience with the product [Venkatesh and

Davis, 1996; Tractinsky, 1997]. User acceptance, or attitude, towards the product, can be described as the user's perceived usability of the product [Vohringer-Kuhnt, 2003]. User acceptance in the TAM literature is therefore belief-based, in that it is based on the users' perceptions of usability. Thus, the main argument against the relationship between user acceptance and usability is based on the differences between perceived usability (belief-based) and actual performance (performance-based).

The definition of usability as stated in the ISO 9241 standard [1997] indicates that usability is both a belief-based and performance-based concept. Effectiveness and efficiency are undoubtedly performance based concepts, in that they are measured in terms of actual performance. Satisfaction is defined by this standard as the comfort and acceptability of the work system to its users. Measures of satisfaction may relate to specific aspects of the system or may be measures of satisfaction with the overall system. According to the customer satisfaction literature, users develop a value judgment about their level of satisfaction with a specific system characteristic. Overall customer satisfaction is then developed based on an aggregate of the satisfaction levels with the various system characteristics [Bietz and Olson, 2001]. This is illustrated in Figure 2.4.

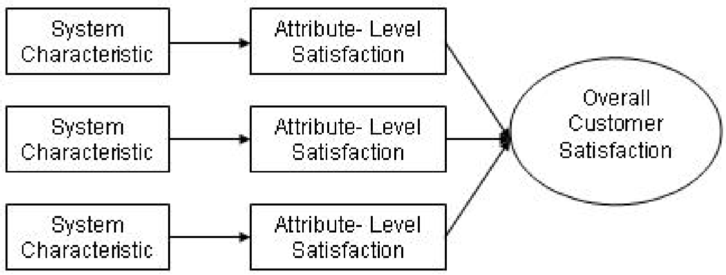

Figure 2.4: Overall Customer Satisfaction [Bietz and Olson, 2001]

Bevan [1995], one of the authors of the ISO 9241 [1997], standard explains that 'measures of satisfaction describe the perceived usability of the overall system by its users and the acceptability of the system to the people who use it and to other people who are affected by its use. Measures of satisfaction can provide a useful indication of the user's perception of usability, even if it is not possible to obtain measures of effectiveness and efficiency.' Satisfaction is therefore a belief-based concept that assesses the users' perceived usability of the system. Therefore, according to the ISO definition [ISO 9241, 1997] of usability, perceived usability is a valid measure of usability.

In addition, the concept of usability has been placed within the framework of system acceptability. Together with other system attributes, such as cost, utility and likeability, system usability determines whether people will accept (purchase and use) a computerized system

[Shackel, in Tractinsky, 1997; Nielsen, 1993]. First impressions often influence acceptability to a large degree [Tractinsky, 1997]. In a study of information systems use, researchers found that 'if computers were perceived initially as difficult to use, users were more likely to express dissatisfaction after four months of use' [Hiltz and Johnson, in Tractinsky, 1997].

The differentiation between performance-based usability and belief-based usability suggests that designers should differentiate between usability goals and user experience goals [Preece et al., 2002]. Differentiation between usability goals and user experience goals allows for the identification of two categories of valid, measurable usability parameters, namely objective usability and subjective usability [Bailey, 1996]. Objective usability measures assess how capable users are at using the system, and include measures related to effectiveness and efficiency. Subjective usability assesses how satisfied the users are with the system, and includes measures of acceptance and comfort. The TAM literature shows that if users are satisfied with the perceived usability of the system, they will accept the system, thus indicating that the system is subjectively usable.

The above discussion has identified the following additional terms and definitions relevant to this research:

- Usability: usability comprises subjective usability (satisfaction) and objective usability (effectiveness and efficiency). Consequently, a relationship between subjective culture and usability can exist as a relationship between subjective culture and objective usability, or as a relationship between subjective culture and subjective usability, or both.
- Acceptance*:* acceptance, as defined by the TAM literature, is a measure of satisfaction and therefore of subjective usability. Therefore, any variable that influences acceptance should be taken into consideration when researching the relationship between culture and usability. Variables that influence user acceptance are discussed in more detail in Chapter 5.

Having shown that user acceptance is related to usability in terms of subjective usability measures, we now present four studies that provide evidence that subjective culture is related to user acceptance. These studies are discussed on a case-by-case basis below.

a. The Straub, Keil and Brenner Study

The Straub et al. study [1997] investigated the possible effects of Hofstede's (1987) cultural model on user acceptance of e-mail technology. Straub et al. [1997] suggested that user acceptance of communications media would be affected by Hofstede's cultural dimensions due to the levels of information richness and social presence afforded by the technology. Media that are information rich and support social presence allow users to experience others as being psychologically present. Socially-present media, thus, are known for their ability to ' transmit information about facial expression, direction of looking, posture, dress and non-verbal cues' [p

11].

Straub et al. [1997] hypothesized that Hofstede's dimensions would affect user acceptance in the following ways:

- Uncertainty avoidance: High uncertainty avoidant users would accept technology that reduces uncertainty and ambiguity by being able to transmit information pertaining to facial expression, direction of looking, posture, dress and non-verbal cues. Therefore, communication technology such as email would be considered unsuitable.
- Power distance: High power distant users would find the levelling affect of technology, such as email, unacceptable. In high power distant societies, individuals deliberately refrain from using email to communicate with their superiors as a mark of respect.
- Individualism vs. collectivism: Collectivist users would probably reject communication media that suppresses the group effect. For example, email does not allow the users to pick up cues about the social situation very easily.
- Masculinity vs. femininity: Masculine users are assertive in nature, and assertiveness requires a social presence. Hence, masculine users would find communication media that does not convey their social presence unable to transmit the overall message of the communicator, and would therefore be unacceptable.

Although Straub et al. did not consider the time orientation dimension, we propose that short-term oriented users would not be comfortable with communication media that caused delays in communication. For example, communication via email requires that the communicator wait for a reply. Short-term oriented users who require quick responses so that they can complete their tasks quickly, would find this waiting period unacceptable.

As a result of the expected influence of these cultural dimensions on user acceptance, Straub et al. developed a Computer-based Media Support Index (CMSI) that 'mathematically expressed the combined effect of the four Hofstede dimensions on the acceptance of e-mail by different cultures'. This index was calculated by adding the scores obtained by each country for each dimension as reported in Hofstede's study. The dominant cultural profiles of each of the countries, as reported by Hofstede [2001], are depicted in Table 2.4, together with the scores obtained.

Country	Power Distance	Uncertainty Avoidance	Masculinity vs. Femininity	Individualism vs. Collectivism	CMSI
USA	40 Low power distant	46 Low uncertainty avoidant	62 Masculine	91 Individualist	157
Switzerland	34 Low power distant	58 Medium	70 Masculine	68 Individualist	204
Japan	54 High power distant	92 High uncertainty avoidant	95 Masculine	46 Collectivist	287

Table 2.4: Cultural Profile, Scores and CMSI of Each Country

Based on the CMSI, the following two hypotheses were formulated:

(1) Users with a high CMSI index, such as Japanese users with a cultural profile of high uncertainty avoidance, high power distance, collectivism and masculinity, would reject communication media that are not information rich or do not support social presence.

(2) Users with a medium to low CMSI index, such as American and Swiss users, would accept such communication media.

The findings of this study supported the hypotheses. This suggests that the users' subjective cultural profile influences their acceptance of particular technologies to perform particular tasks. This once again provides support for the theory that there is a relationship between subjective culture and usability.

b. The Massey, Hung, Montoya-Weiss and Ramesh Study

The Massey et al. [2001] study investigated the possible effects of uncertainty avoidance, individualism / collectivism and context on cross-cultural communications in global virtual teams. They hypothesized that these three subjective cultural dimensions in conjunction with each other reflect the effects of culture on such communication.

Massey et al. [2001] hypothesized that these three dimensions would affect the users' perceptions of the 'fit' between the technology and the communication tasks. The results of their study supported their hypothesis in that the users' perceptions of the task-technology fit were mediated by the user's cultural profile. Thus, cultural differences in terms of uncertainty avoidance, context and individualism/collectivism affect communication behaviour and hence perceptions and preferences. This once again suggests that subjective culture influences user acceptance of particular technologies to perform particular tasks, providing more support for the theory that there is a relationship between subjective culture and usability.

c. The Smith and Chang Study

Smith and Chang [2003] conducted a study to determine the extent to which Hofstede's [2001] cultural dimensions influenced Chinese users' acceptance of interfaces. Their study reported that Chinese users preferred sites that displayed high power distance, high uncertainty avoidance, masculinity and individualism characteristics. It was found that the contributions to variance by these dimensions were different. Power distance contributed 22% to the variance in user preference, whilst masculinity vs. femininity and individualism vs. collectivism had lower contributions of 9% and 2% respectively, and uncertainty avoidance virtually no contribution at 0.01%.

The findings of this study provide empirical evidence that subjective culture influences users' preferences and acceptance of website interfaces, thus suggesting that there is a relationship between subjective culture and usability. Interestingly, these findings also suggest that only some of the cultural dimensions have a significant impact on usability. This also highlights the possibility that the different cultural dimensions may have a stronger or weaker effect on usability. These issues will be discussed in more detail in Chapter 5.

d. The Anandarajen, Igbaria, and Anakwe Study

Anandarajen et al. [2002] studied the extent to which Hofstede's [1991] cultural dimensions influenced user acceptance of technology, specifically in the context of a collectivist society. They also studied the effects of Hofstede's dimensions on selected variables that have been proposed to influence user acceptance, including computer playfulness, organizational support and individually-based computer training. They hypothesized that:

(1) Individuals from a collectivist culture would use computers not because of their perceived usefulness or the enjoyment to be derived, but rather because of the perceived social pressure from their peers.
(2) Computer playfulness would not influence acceptance for high uncertainty avoidant users, as playfulness involves creativity as well as unstructured experimentation.
(3) Organizational support would influence acceptance for collectivist and feminine users, as these societies value cooperation and harmony.

The findings of this study supported the above hypotheses, which raise doubts as to the applicability of the Technology Acceptance Model for cultures that are different to the dominant culture of North America. This study reported that cultural tendencies towards high uncertainty avoidance, high power distance, collectivism and masculinity disassociate usefulness from acceptance. In addition, perceived enjoyment was found to have no direct effect on acceptance, and individual training did not enhance computer skills and therefore had no effect on perceived ease of use. In contrast, it was found that organizational support was positively related to perceived ease of use in a collectivist society. Thus, the findings not only confirm that there is a relationship between subjective culture and usability, but that the effects that related variables have on user acceptance is dependent on the subjective cultural context of the intended users.

The four studies discussed above provide strong theoretical evidence that subjective culture is related to user acceptance. As user acceptance is a valid measure of usability, this provides further support that a relationship between subjective culture and usability exists from a user acceptance perspective. We now turn our attention to the objective usability perspective.

2.6.2.2 The Objective Usability Perspective

As discussed in the previous section, objective usability assesses how capable users are at using the system, and include measures related to the accuracy, completeness and speed with which users complete specific tasks. Empirical evidence that subjective culture influences objective usability was found in the study performed by Forer and Ford [2003]. This study investigated the relative effects of heuristics and subjective culture on usability. It was hypothesized that:

(1) Greater user performance will be achieved on an interface designed to match the users' cultural profile in terms of Hofstede's [2001] cultural dimensions, than will be achieved on an interface designed to only accommodate Nielsen's [1993] usability heuristics.

(2) Greater user performance will be achieved on an interface designed to accommodate both Nielsen's usability heuristics and the user's cultural profile in terms of Hofstede's [2001] cultural dimensions, than will be achieved on an interface designed to only accommodate for Nielsen's [1993] usability heuristics.

(3) Greater user performance will be achieved on an interface designed to accommodate both Nielsen's [1993] usability heuristics and the user's cultural profile in terms of Hofstede's [2001] cultural dimensions, than will be achieved on an interface designed to only accommodate the users' cultural profile in terms of Hofstede's cultural dimensions

The study reported significant differences in accuracy and speed between the interfaces. No significant differences were found for overall satisfaction between the interfaces; however, different satisfaction ratings were found between the appearance, metaphors and interaction components of the different test interfaces. It was concluded that the use of an interface that displays a cultural profile matching the user results in better usability than the use of an interface that displays an opposing cultural profile, irrespective of whether or not Nielsen's [1993] usability heuristics are accommodated. This provides some empirical evidence that there is a relationship between subjective culture and objective usability, for at least one cultural profile.

We will now look at the relationship between subjective culture and usability from a context of use perspective.

2.6.2.3 The Context of Use Perspective

As discussed in section 2.4.2, the context of use has been described in terms of the three contexts of the user, the task and the environment. Theoretical support for the inclusion of culture as a dimension of the user context was found in two literature sources, which are discussed next.

Firstly, the Hall, Lawson and Minocha study [2003] suggests that subjective culture provides the context within which different solutions to the same problem can be identified. Whilst the problems faced by different cultures are the same, the way in which those problems are solved is dependent on the cultural context of the users. For example, a common problem faced by designers of web interfaces is to find a way that will attract users to the site. For individualist cultures, one way to solve this problem would be to focus on the goals of the individual. In contrast, focusing on collective agendas would be a more suitable solution to the same problem for collectivist societies [Marcus, 2000]. This problem-solving context can be extended to the other cultural dimensions identified by Hofstede as well. This suggests that subjective culture can be seen to provide the context within which the problem should be solved. This again indicates a relationship between subjective culture and usability, as the context of use influences the usability of the product. An example of the power distance cultural dimension is illustrated in Table 2.5.

PROBLEM	ATTRACT USERS TO THE WEBSITE				
Dimension	**Individualism vs. Collectivism**	**Power Distance**	**Uncertainty Avoidance**	**Masculinity vs. Femininity**	**Time Orientation**
Context	Individualist	High power distant	High uncertainty avoidant	Masculine	Short-term oriented
Solution	Focus on visitor goals	Provide complex, highly organized structures	Provide limited choices and binary logic	Focus on work goals	Focus on liberty, social incoherence and efficiency
Context	Collectivist	Low power distant	Low uncertainty avoidant	Feminine	Long-term oriented
Solution	Focus on collective agendas	Provide simple, informally organized structures	Provide multiple options and fuzzy logic	Focus on relationships	Focus on social responsibility, coherence and support

Table 2.5: Subjective Culture as a Problem-Solving Context

Secondly, in a panel discussion held during the 2nd HCI and Culture Workshop, it was reported [Light, 2003] that Gould pointed out that subjective cultural dimensions 'serve as design opportunities'. For example, high power distant users that have a high respect for superiors would appreciate phones that 'know when to interrupt you'. Similarly, phones that provide for asynchronous interactions would be more suitable for monochronic users than for polychronic users; games and one to one calls would be appropriate for individualists who put their own interests first; whilst phone features such as buddy lists and emergency phone trees would be more suitable for collectivists. This highlights that subjective culture influences the task context of use, as users with different cultures will want to perform different tasks appropriate to their respective cultural traits. This once again suggests that subjective culture influences the context of use.

2.6.3 Operationalisation of Proposed Metamodel

We have shown that Hofstede's cultural dimensions theoretically influence usability in terms of user acceptance, context of use and objective usability. Consequently, we can operationalise the metamodel proposed in section 2.5.5 by providing an example of how this metamodel is applied using Hofstede's model of culture. Using the cultural user interface design guidelines proposed by Marcus [2001] for high power distance as an example, this approach would result in the use of the interface design guidelines illustrated in Table 2.6.

Subjective Culture ↓	**SUBJECTIVE CULTURE** **High Power Distance**				
Subjective Interface Design Guidelines ↓	Buildings with clear hierarchy	Reference data with no relevance rankings	Restricted access and choices	Images of leaders, official music and formal speech	Severe error messages and the use of wizards
User Interface Components	**Metaphors**	**Conceptual Models**	**Navigation**	**Appearance**	**Interaction**
↑ *Objective Cultural Design Guidelines*	Buildings relevant to country or region	Sort data according to occupational requirements	Restrict access to non-members	Images of leaders relevant to religion	Language relevant to linguistic affiliation
↑ *Objective Culture Levels*	**Country**	**Profession**	**Social Class**	**Religion**	**Language**
↑ *Objective Culture*	**OBJECTIVE CULTURE**				

Table 2.6: Example of Proposed Approach for High Power Distance

However, there is no empirical evidence that Hofstede's cultural dimensions are actually applicable to, and significant within, interface usability, and in particular, website usability [Smith et al., 2004]. Consequently, we tested our theory by performing an experiment that tested the effects of these dimensions on user performance. The experiment is discussed in detail in the next chapter.

2.7 Summary

This chapter presented a review of the literature focusing on the influence of culture on usability and performance. We began by discussing human-computer interaction, which forms the context within which this research is based. This was followed by a discussion of human performance, which we argued to be the primary objective of the study of human-computer interaction. The cognitive processes involved in human performance were reviewed, followed by an analysis of how factors such as the human information processing system and user characteristics influence the cognitive load and resultant performance.

Thereafter, the concept of usability was discussed from the perspective of enhancing user performance. We showed that usability is not a one-dimensional concept, but rather consists of multiple characteristics that should be considered when evaluating user interfaces. We pointed out that usability is a function of the context of use, and therefore the usability characteristics that are important are dependent on the user, the task and the environment within which the interface is being used. This led to the conclusion that designers are designing the relationship between the product and the human, not the product per se [Rubin, 1994; Gould et al., 2000]. We identified numerous usability principles that have been proposed as a way to enhance this relationship, and argued that compatibility, rather than consistency would improve performance if the user interface is designed to be compatible to the contexts within which the system is used.

Our attention then turned to culture, which can be seen to be part of the context of use. We proposed that in order to design systems that are fully compatible to the context of use, it is necessary to accommodate cultural diversity into the user interface. Four metamodels and four models of culture were reviewed, and the arguments for and against the importance of culture in human-computer interaction were summarised. It was argued that culture influences communication, and as interactive systems are dependent on communication, cultural diversity should be accommodated into user interface design. We then reviewed two approaches to culturalisation: the first approach is based on translating only the objective cultural aspects of the interface; the second approach is based on the premise that subjective culture, although perhaps more difficult to manage, should also be included. We argued that both objective and subjective culture should be incorporated, and proposed a metamodel for accommodating subjective cultural aspects using a top-down approach, and objective cultural aspects using a bottom-up approach.

The metamodel that we proposed highlighted the need for relevant guidelines for accommodating culture into the design of interfaces. Four models that have previously been used as a basis for such guidelines were reviewed. We discussed the advantages and pitfalls of each, noting that the cultural dimensions model, and in particular Hofstede's model of culture, has been criticized in the literature. However, by synthesizing relevant studies from the literature, we provided theoretical evidence that the cultural dimensions comprising Hofstede's model influence the cognitive load, thus affecting performance. Further; evidence that these dimensions are related to usability from the user acceptance, context of use and objective usability perspectives was presented. Consequently we concluded that using Hofstede's model is a relevant basis for managing the subjective cultural aspects of interface design. Using one of Hofstede's dimensions, we provided an example of how to operationalise the metamodel that we proposed.

Whilst there is sufficient theoretical evidence to suggest that Hofstede's cultural dimensions influence usability and performance, there is no empirical evidence to support this. Consequently, we conducted an experiment to test our theory by testing the effects of these dimensions on user performance. We provide an overview of the research design and methodology used in the experiment in Chapter 3, followed by a detailed discussion of the experiment and the results in Chapter 4.