First Impressions Of Operating System Styles Affect Usability

Prajesh Narendra Chhanabhai

A dissertation submitted for the partial fulfilment of the requirements for the degree of Masters of Science at the University of Otago, Dunedin, New Zealand

12 November 2004
“Even as wisdom often comes from the mouths of babes, so does it often come from the mouths of old people. The golden rule is to test everything in the light of reason and experience, no matter from where it comes.”
Mahatma Gandhi
PART A

RESEARCH REPORT
Abstract

The patterns of behaviour that people develop to work successfully with complex information technology are likely themselves to be complex. The beginning of interface style followed this train of thought in developing the Command Line Interface, complex to develop and complex in its use. However, in recent years' information system interface design has become increasingly dominated by the use of Graphical User Interfaces, with the majority of systems relying on a Microsoft Windows based structure. This study attempts to find out how the two different interface styles, would affect how novice users use them when given a word processing task. The study was conducted using two interfaces based on a graphical style, and two command line type operating systems. The study examined how quickly the participants performed the task on the different interface styles, and used questionnaires to gather the quantitative findings. Preference versus performance was studied and the findings are consistent with what other researchers have found. The results indicate that better usability may not mean better performance.
Acknowledgements

The field of research is exciting, hard work and always rewarding. But as a first time researcher this field is full of daunting yet novel possibilities. It is in this regard that I would like to thank a few people that have helped me through this experience.

Samuel Moyle, or Sam as he likes to be known. My Supervisor, thank you for all the help, advice and direction you gave me. Thank you for labouring over my draft copies of various pieces of paper and proof reading my work. Your knowledge in the field of Human Computer Interaction proved invaluable.

The “Crew” at TSG for giving me machines to play around and then run my experiments on. A particular mention must go to Brendon Sly for his patience in helping me get the machines, especially the Macintosh, to run as we wanted them to.

Statistics not my forte, but thanks to Brian Niven in the Statistics Department, for taking time out of his busy schedule to go through my SPSS data sheets and help me understand what all the numbers mean.

To Annie and Josef, thank you for advising me on a number of things that I would not have known about if it hadn’t been for your guy’s information. And for the attempts at trying to explain how SPSS works.

Thanks must go to someone who was always there for me, and always pushing me on, asking about how my work was going and making sure I kept at it, Darshana. Thank you for your time and for being understanding throughout the research process.

To my fellow peers from the Info 480 class, for always being on hand to chat to about my project and sharing this “ordeal” through the good and bad. A special mention must go to Matt and to Dennis who were always up at odd hours in the MSRL working away, making the task of writing slightly easier as there was always time for some humour and chat, to relax the nerves.

Lastly but by no means the least, I’d like to thank my family, who are on the other side of the world. For giving me the opportunity to come and study in New Zealand and allow me to experience research first hand. Even though you are in Zimbabwe, your words of advice always came in handy.

Thank you.

“We do not believe in ourselves unless someone reveals that deep inside us is valuable, worth listening to. We tell others of our faults, guard in our hearts. Once we believe in ourselves we can risk curiosity, wonder, spontaneity, delight or any experience that reveals the human spirit.” - E.E. Cummings
# Contents

1. INTRODUCTION  
   1.1 General Background for the Study ........................................ 1  
   1.2 Definition of Terms .......................................................... 2  
   1.3 Research Problem .............................................................. 3  
   1.4 Hypothesis ............................................................................ 3  
   1.5 Variables .............................................................................. 3  
      1.5.1 Independent Variables .................................................... 3  
      1.5.2 Dependent Variables ......................................................... 3  
   1.6 Assumptions ......................................................................... 4  

2. REVIEW OF LITERATURE  
   2.1 Difference Between A Novice and An Expert ............................ 13  
   2.2 Previous Experimental Methods ............................................. 16  
   2.3 Summary ............................................................................... 19  

3. METHODOLOGY  
   3.1 Sample ................................................................................ 20  
   3.2 Instrumentation .................................................................... 20  
   3.3 Data Collection ..................................................................... 21  
   3.4 Justification of Questions ....................................................... 23  
      3.4.1 Pre-Observation Questionnaire ........................................ 23  
      3.4.2 Post-Observation Questionnaire ....................................... 24  
   3.5 Data Analysis ........................................................................ 26  

4. RESULTS  
   4.1 Numerical Findings ............................................................... 27  
   4.2 Statistical Findings ................................................................. 28  
      4.2.1 Time Taken .................................................................... 28  
      4.2.2 Time Perception ................................................................ 30  
      4.2.3 Perceived Expertise .......................................................... 33  
      4.2.4 Satisfaction ................................................................. 34  
   4.3 Summary ............................................................................... 38  

5. DISCUSSION  

6. CONCLUSION  
   6.1 Confounding Factors ............................................................ 44  
   6.2 Future Work ....................................................................... 46  

REFERENCES  

APPENDICES  
   Appendix A  Information Sheet and Consent Form ......................... 52  
   Appendix B  Pre-Observation Questionnaire .................................. 56  
   Appendix C  Task Sheets ............................................................ 56  
   Appendix D  Post-Observation Questionnaire ................................ 60
List of Tables

Table 1 Comparisons between Human and Computer Performance .................. 5
Table 2 Interface requirements ......................................................................... 8
Table 3 Times on the four interfaces ................................................................. 27
Table 4 Number of Participants vs. Time Perception ........................................ 31
Table 5 Maximum and Minimum Times on the Interfaces ................................ 31
Table 6 Mean Expertise GUI vs. CLI ................................................................. 34
Table 7 Mean Expertise vs. Time Perception .................................................... 34
Table 8 Participants responses to satisfaction questions .................................. 35
List of Graphs and Figures

Graphs
Graph1 Mean total time against the type of interface used .......................... 28
Graph 2 Time taken on the two interface styles ........................................ 30
Graph 3 Perception of time taken for participants on GUI and CLI ............... 32
Graph 4 Correlation between the logs of actual time against perceived time taken .......................... 32
Graph 5 Perceived expertise of use on a GUI and a CLI .............................. 33
Graph 6 Percentage of users who found the Interface styles satisfying ........... 35
Graph 7 Percentage of users who did not find the Interface styles satisfying ..... 35
Graph 8 Number of participants and their answer towards the satisfaction questions against their perceived expertise ............................................................... 37
Graph 9 Number of participants and their answer towards the satisfaction questions against their perceived expertise ............................................................... 37

Figures
Figure 1 A classification schema of most common user interfaces ................. 7
Figure 2 A space of interfaces ..................................................................... 12
Figure 3 Comparison of Inter and Intra Interface comparisons ................... 21
Figure 4 Distribution of the interfaces with regards to time ....................... 29
1. Introduction

So many operating systems and so many graphical desktop environments... System developers face many challenges and decisions as they analyse and create designs for the various types of interfaces around. One of the biggest factors for them to consider is how the system will function and operate for the intended user (Merchant, 2002). Jefferson (2002) postulates that the two main reasons for continued “lousy interfaces” are the reduction in the importance of end-user experience and design based on the requirements of operating systems rather than those of the end-user. Manes (2000) suggests that there is a growing public sentiment of interface dissatisfaction as most of the interface styles anger and confuse users.

There have been numerous studies carried out over the years, where researchers have pitted the Command Line Interface (CLI) against the Graphical User Interface (GUI). Many interesting findings have been uncovered. This study will look into some of the research that has occurred in this field and extract the essential findings from these studies. The structure of this paper begins with a short background, followed by a review of the existing literature, tackling various issues that have been addressed in previous studies. Methodology and results analysis will follow the literature review and the paper will end with a discussion and conclusion.

1.1 General Background for the Study

Information technology (IT) has seen billions of dollars being invested into the sector, however, studies have shown that white collar productivity is still low (Smith, 1998). The relationship between IT investment and productivity is alarming to organisations and individuals alike. Smith (1998) concludes that considerable time and money is being spent to acquire, train users and use such technology. The key barrier to the low productivity is that many users of IT do not actually know how to use it and many people are scared of using such technology. Some contributing factors are the inherent usability of the system (Schneiderman, 1982a), functionality it provides (Davies, 1989) and the man-made interface (Schneiderman, 1982b). It can be argued that of the three factors the one that affects IT users most is the man-made interface, as its design determines eventual user acceptance and utilisation (Davis and Bostrom, 1993). Graphical User Interfaces (GUIs) are on their way to becoming the most pervasive interface for desktop systems at least partly because of assumed ease of use. Such an assumption may have been kindled by vendors' claims about the inherent
First Impressions of Operating System Styles Affect Usability

usability of such interfaces although previous research on the productivity gains from GUIs has yielded mixed results (Zanino et al, 2001).

The purpose of this study is to assess the usability of four operating system styles (user interfaces) that are used today. As it can be seen the problem found is that there is low productivity, even through the use of IT. Type of interface may be linked to productivity, but this study will look more into the relationship between performance and preference which may infer a link to overall productivity.

1.2 Definition of Terms

The paper will have a number of terms that will be used frequently. The key terms that are directly related to the research will be defined now. This will ensure the reader has the same idea as the researcher, when they come across these terms.

**Human Computer Interaction (HCI):** 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 (ACM SIGCHI, 2004)

**Operating System:** The most important program that runs on a computer: the Operating System (OS) performs basic tasks such as recognising input from the keyboard, sending output to the display screen, keeping track of files and directories on the disk, and controlling peripheral devices such as disk drives and printers. Operating Systems include DOS, Windows, OS 9.2, OS X, Linux or UNIX. (Glossary Online, 2004)

**Interface:** In relation to human communication with a computer, the appearance of the screen via which the interaction occurs. (Glossary Online, 2004)

**Novice:** Any user, unfamiliar with computer use. For our purposes, it is taken that students currently enrolled in COMP101 at the University of Otago to be novice users.

**Command Line Interface (CLI):** An interactive user interface which allows commands to be given to a computer program or shell through a text-based terminal. (Approved LPI Glossary, 2000)
1.3 Research Problem

To what extent is usability affected by the type of interface used, when presented to a group of novice computer users?

1.4 Hypothesis

It is hypothesised that providing novice users with certain operating system styles (GUI and CLI) will influence their perception of computer usability. Novice users are expected to prefer and be more productive using a GUI interface. The hypothesis can be broken into two sections.

1. Novice Users will prefer a GUI to a CLI.
2. Novice users will be more productive using a GUI than a CLI.

1.5 Variables

The variables that will be investigated in this study are

1.5.1 Independent Variables

This will be an active independent variable as the researcher will be able to monitor its effect on the dependent variables. The independent variables in this study are the Operating Systems or Environments. This can be further classified as four independent variables GUI preference, GUI performance, CLI preference and CLI performance.

1.5.2 Dependent Variables

The key dependent variables under study will be the usability, impressions and quality of human computer interaction. These will be obtained by looking at perceived time to perform a task, actual time to perform a task, perceived expertise on the interface style used. Essentially indicating ease of use, speed and perceptions.
1.6 Assumptions

The following assumptions have been made:

- The users will be able to perform the tasks given to them.
- Users will have limited interaction with the different interfaces before this study is carried out.
- The tasks that will be given will be ones that non-IT users of computers will be likely to carry out on a day to day basis.
- The study environment should not affect the outcomes of the study.
- The novice users to which the tasks will be given are indeed novice users.
2. Review of Literature

If you asked most people in the world today how easy they found a spoon to use they might wonder why you asked. Such things are so well designed for their purpose that we take them for granted — it is only when objects are hard to use or lead to errors that one tends to focus on them (Lansdale and Ormrod, 1994). Some of these breakdowns are physical or physiological in origin: “an information display might be obscured by a glare” (Cushman and Crist, 1987, p. 271) or a particular keyboard configuration might cause some form of physical strain (Grandjean, 1987). As creatures lacking in power, speed or the ability to enable rapid turnover humans have relied for survival upon the ability to learn and adapt. People turn to computers for many reasons (see Table 1), to try make up for their inadequacies.

<table>
<thead>
<tr>
<th>Facility</th>
<th>Human</th>
<th>Computer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculation</td>
<td>Slow, error-prone</td>
<td>Fast, accurate</td>
</tr>
<tr>
<td>Vigilance</td>
<td>Intermittent, prone to fatigue</td>
<td>Permanent</td>
</tr>
<tr>
<td>Judgement</td>
<td>Biased</td>
<td>Predictable</td>
</tr>
</tbody>
</table>

Table 1. Comparisons between Human and Computer Performance (Lansdale et al, 1994).

In addition to the factors mentioned in the table, other aspects have also led to the increase of computer use. Boredom is one of those factors; humans are creatures that bore easily when they do a task that is monotonous or one that does not spark interest, a problem not encountered by computers doing the same tasks.

The usability of menu interfaces compared to command line interfaces began to appear in literature after the introduction of direct manipulation input devices, such as the mouse (English at al, 1967). The purpose of these direct manipulation devices was to simplify human-computer interactions (Perry and Voelcker, 1989; Shneiderman, 1983). It was considered that menu interfaces required a less of a cognitive effort than that of CLI's. However despite that assumption, Liu (1997) and Shneiderman (1998) comment that novice computer users are seen to prefer using a menu driven interface whilst expert users prefer to
use the power of the command line. That conclusion is based upon the assumption that recognising a correct command is easier than recalling a correct command. Paap and Roske-Holstrand (1988) acknowledge that the usability advantage of one interface over another is influenced by many interacting factors, including task complexity, user objectives, user facility in manipulating different input devices, and user knowledge of a particular application domain.

Studies that have assessed interfaces for the effectiveness of learning a new skill or for solving a problem have typically compared icon selections with a mouse against text input via a keyboard (Davis and Bostrom, 1993; Schur, 1996).

Computer usage has increased dramatically. In the USA during the 1980’s, the ratio of office workers to display terminals was on the order of 10 to 1 (Galitz, 1993). With the increase in availability of computers this ratio has dropped to almost 1 to 1. This has called for a refocusing of system design to the user. The ease and effectiveness of human interaction with computers depends on how well the interface style reflects user needs (Dunsmore, 1982).

Aprile (2004) brings up an interesting point saying that human beings are not brought up to intuitively handle the GUIs of modern day computing systems. To master modern GUIs, one must recall the operation, layout and relation to each other of hundreds, if not thousands, of such panels. This requires not only time, but some understanding of how each operation may affect the final outcome of the intended product or service. The hardest of these skills is the ability to recall or discover the correct sequence of operations on one control panel relating to a desired operation (Aprile, 2004).

After direct binary switches (which was the way interaction with computers was carried out) humans began using CLI more frequently, the main reason being that it was a more usable interface than the binary switch. Computers during the 1970’s and early 80’s were seen to be for experts at the systems and not for everyday users. Eberts (1994) brought forward the concept that CLI’s, which come with commands and keystrokes that a user must memorise in order to do even basic tasks is something that an everyday user would not bother to take time to try to master.

CLI’s were still in use after the introduction of the more visual interfaces, which arrived in about 1984. The reason for this is given by Raymond (2003), who states that CLI’s “were and are still seen to be more expressive than visual interfaces, for complex tasks” (Raymond,
2003, p. 46). The GUI became popularised in 1984 with the commercial release of the Apple Macintosh. The Macintosh’s arrival was shortly followed up by Microsoft releasing a GUI for the IBM PC. The GUI has become the most popular interface, being used on almost 30 million computers worldwide (Luhman, 1999).

According to Rauterburg (1996) the classification of interfaces is important; two quantitative measures that are suitable to allow this classification are functional feedback and interactive directness. Using these it is possible to classify the most common interface types: batch, command, menu and desktop. The command language interface is characterised by high interactive directness, but this interface type has a very low amount of visual feedback. Only GUIs are found to support the user with sufficient visual feedback and with high interactive directness. To make this classification as understandable as possible, Rauterburg (1996) describes the four classified interfaces with an abstract schema (see Fig 1). The schema includes the batch interface as well, which is not a part of this study. The schema also classifies menu and direct manipulation devices as two separate interfaces, according to Schar (1996) menus are a type of direct manipulation interface hence could be classed as one.

![Figure 1. A classification schema of most common user interfaces. (Rauterburg, 1996.)](image)

Rauterburg (1998) has also observed the trend in what people are seeking when it comes to interfaces styles, and has tabulated opinions of what a good interface should provide. As can be seen in Table 2, the requirements from 1988 till 1996 have stayed relatively similar; however the documents that Rauterburg (1998) acknowledged identified different methods to
quantify similar attributes. This could be attributed to the fact that each has taken interface requirements from a different angle. The increase in the use and implementation of different styles of interfaces, especially the GUI, would have a major influence on these studies, as the GUI has evolved as time has progressed as well as the way people use software.

Table 2. Interface requirements (Rauterburg, 1998)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Suitable for the task</td>
<td>Suitable (activity adapted)</td>
<td>Self-descriptiveness</td>
<td>Task orientation</td>
</tr>
<tr>
<td>Self-descriptiveness</td>
<td>Feedback about the system states</td>
<td></td>
<td>Transparency</td>
</tr>
<tr>
<td>Conformity with user expectations</td>
<td>Conformity with user expectations</td>
<td></td>
<td>Compatibility</td>
</tr>
<tr>
<td>Information and Instruction of User</td>
<td>Suitability for Learning</td>
<td></td>
<td>Support</td>
</tr>
<tr>
<td>Ease of use applicable to skill level</td>
<td>Suitability for Individualisation</td>
<td></td>
<td>Selection possibilities</td>
</tr>
<tr>
<td>Hearing and participation of users</td>
<td>Controllability</td>
<td></td>
<td>user definability</td>
</tr>
<tr>
<td>Controllability</td>
<td></td>
<td>Controllability</td>
<td></td>
</tr>
<tr>
<td>Error Robustness</td>
<td></td>
<td>Error Tolerance</td>
<td></td>
</tr>
</tbody>
</table>
The GUI interface is one that is built using graphical designs and allows the ability to accept input from sources other than the keyboard, eliminating the need to memorise complicated keystrokes (Luhman, 1999). Luhman (1999) suggests that the reason first time users of computers would prefer a GUI is because it provides what-you-see-is-what-you-get (WYSIWYG). Raymond (2003) claims that the biggest factor in the success of GUI’s is that they are designed using metaphors, the most well known being the desktop metaphor. The fact that GUI’s have attempted to standardise entire systems of applications, and that programs are all arranged in a similar manner means that the learning curve is greatly reduced (Luhman, 1999).

One important method of assessing how good interfaces are is to measure their usability. Nielsen and Levy (1994) suggest that measuring usability falls into two broad categories:

1. Subjective user preferences which measures how much the users like the system.
2. Objective performance measures which measures how capable the users are at using the system.

One would expect that there would be a direct positive correlation between subjective preference and objective performance, as people can be expected to prefer doing their job well and to use computers with interfaces that help them rather than hinder them (Nielsen and Levy, 1994). This is true for most cases; however there are examples when users do not like using a system as it slows down their objectives, even though they like its overall layout. For example MacLean et al (1985) found that subjects preferred the slower of two data entry methods as long as it was not 20% slower than the faster method.

Usability includes the ability for a computer user to find what they are looking for with relative ease. This is especially true when they require help. Most interfaces these days come with some sort of a help option, however, Wright (1983) finds that people that work with a variety of tools have been generally found that instead of reading the documentation to find out how to do something, they will “play with the instrument itself”, or ask “someone who knows”. She also reports that studies of the use of manuals by novices indicate that the majority prefer to leaf through a document as opposed to using contents lists or index. She also indicates that “many of the people who turn to a manual will be looking for answers to questions about what to do, rather than seeking an understanding of why they should do it” (Wright, 1983, p. 2).

Wright (1983) suggests that people who read manuals draw expectancies of the manual, based on past experience. Readers sometimes draw “inferences from the text which are not
intended" (Wright, 1983, p. 13) by the authors, resulting in unexpected consequences leading to the reader having to conduct further research to carry out their initial task correctly (Wright, 1983, p. 13).

Shneiderman (1982) has drawn up a list of attributes, that an interface should meet, if it is to be regarded as a usable interface:

1) Learnability
2) Enhanced expert performance
3) Fewer Error Messages
4) Better Feedback
5) Memorability
6) Reduced anxiety
7) Increased Control

Shneiderman (1982), regards most GUI’s as conforming to these 7 requirements, he however maintains that a GUI alternative does not necessarily make a system more usable than a CLI. The representation must be meaningful and accurate.

GUI’s are thought to be a powerful way of solving many problems associated with computer use as they are a good way to reduce the chance of miscommunication. Often numeric and verbal cues are insufficient to trigger insight into the specific nature of a problem (Larkin and Simon, 1987). GUI’s are used for data reduction, data summary, to improve information search and to facilitate computation (Robertson et al, 1993). Graphics, pictures and diagrams transcend language barriers, an influential factor in a world now known as a global village due to the growth of the Internet.

Cognitive psychologists have found that visual information can benefit behaviours such as recall, comprehension and retention of information (Umanath and Scamell, 1988). Good use of this knowledge is made use via the ‘desktop’ metaphor. This metaphor portrays the operating system of the computer as similar to objects, tasks and behaviours found in a physical office environment, unlike a CLI whose style is a signature for “only those who know computers” (Mynatt and Edwards, 1995, p. 202). Hutchins et al (1986), argue that the users prefer an interface that is direct. One achieves directness through two aspects: engagement and distance.

\[ \text{DIRECTNESS} = \text{ENGAGEMENT} + \text{DISTANCE} \]

(Hutchins et al, 1986)

Engagement refers to the perceived locus of control of action within a system. This shows whether the interface that is being used allows the user to feel that they are the principle actors
First Impressions of Operating System Styles Affect Usability

P Chhanabhai

in the running of the system, or not. In a CLI style of interaction the locus of control appears to reside with a “hidden intermediary” who executes English like commands such as 
\texttt{C:>dele te *.exe}, on the users behalf (Frohlich, 1997). This type of system is indirect. This claim appears to contradict the schema proposed by Rauterburg (1996), in which a CLI is claimed to be regarded as a High level of directness. However, despite the high level of directness, Rauterburg’s (1996) schema indicates that CLI’s have low visual feedback which linked to engagement. In a graphical based style of interaction, the locus of control appears to reside with users themselves who influence the objects of interest within a model world (Frohlich, 1997). Distance is the mental effort required to translate goals into actions at the interface and then evaluate their effects (Norman, 1986). Mental calculations have to be made by users when using various systems, these calculations have a two fold effect. They enable the user to decide what they should do and they allow the user to decide how best to bridge the gap that may exist if they are in an unknown environment. Interfaces that make it easier for users to make these calculations are said to be more direct to use (Norman, 1986). In simple terms this implies that users of GUI’s generally don’t have to “think so hard”, or learn as much when compared to learning and using a command language for the same operations. 

Figure 2 shows that the space of interfaces has an effect on the user. Based on Hutchins et al (1986) the figure shows that the least direct interface is one that provides a low level language interface, as it provides the weakest semantic match between the intentions of the user and the language of the interface. However if the interface is one that is modelled according to the world, i.e. using metaphors, then the user is believed to be directly engaged with the objects and make the process somewhat easier. Sometimes these interfaces look as if they are easy to use but as a user gets to know them better, they actually find the interface is actually quite clumsy.
These are interesting ways to determine the directness of a system, however, both the defined aspects are almost impossible to measure and quantify (Hutchins et al. 1986). It is a theory that involves measures that are wholly subjective, thus making them difficult for one to analyse scientifically. A theory proposed by Hayes and Broadbent (1988), cited in Svendsen (1991), can be useful when comparing interface style on problem solving. The theory declares that there are two types of learning that takes place when doing interactive tasks.

1) U-Mode: where learning takes place outside the working memory, and not verbally reportable.

2) S-Mode: learning takes place in the working memory and can have verbal reports.

The results from Svendsen’s (1991) study bring him to the conclusion that CLI’s induce S-Mode learning whilst GUI’s promote U-Mode learning. They suggest another way of stating the theory mentioned by Hutchins et al (1986), asserting that directness is felt when interaction occurs without the user having to encompass any conscious thought.

Comparisons of the use of menu interfaces and command line interfaces are complicated by the potential structure of a menu interface and by the number of options presented to a user. According to Durham and Emurian (1990) in a menu interface, recognition could occur within a linear or hierarchical format of selectable options. Items to be scanned could appear within the same scroll window, or they could be displayed simultaneously in a separate
window. Menu-based interfaces are sometimes preferred by unskilled users and command-line interfaces are sometimes preferred by skilled users (Joseph et al., 1989) even though, according to Rosson (1983), the menu structure may hamper the learning and performance process. This finding is supported by further observation by Santhanam and Wiedenbeck (1993) who noted that novice computer users, who are considered to be infrequent and not highly skilled or motivated users, often perform better on a menu interface in comparison to a command line interface for a specific application. Other studies (Hauptmann and Green, 1983; Lee et al, 1986), advocate novice users find making selections from a menu slightly more confusing than entering commands on a CLI. Despite the view among interface designers that recognition using a GUI is easier than recall using a command line (Galitz, 1993), studies in the literature document that this view is found to be the exception rather than the norm (Greene et al., 1992). Svendsen (1991) declares that the reason people associate a user-friendly interface as one that makes it a good problem-solving tool is because people assume that:

1) using an interface and solving problems compete for the same cognitive resources; and
2) a user friendly system somehow facilitates problem solving.

According to Svendsen (1991), both assumptions are hard to measure and are not warrantable as many factors affect the user friendliness of a program which will affect problem solving skills as well.

2.1 Difference between a Novice and an Expert

This study is going to look at how novices react to the different interface styles i.e. GUI and CLI on Windows and Macintosh OS’s. Research has been carried out on the difference between novices and experts, this section will elaborate on some of those findings. The definition that was given earlier will be used to describe what a novice user will be regarded as in the context of this research i.e. unfamiliar with computer use. According to Santhanam and Wiedenbeck (1993), novices are people who strive to make sense out of experiences, hypotheses and creating analogies to explain the behaviour of a system. On the other hand the definition of an expert in research literature seems to be related to the purpose and the nature of the research.

When a novice ventures into a new domain such as using a word processing system, a computerised data retrieval system, an operating system or a general purpose programming language, it can be observed that the user makes many errors and requires a lot more time to solve problems. However over the course of time and through continuous use the user
becomes faster and more accurate. This transition from a novice to a user that is able to perform functions quicker and with accuracy maybe said to be a transition from being a novice to becoming an expert user. According to Kolodner (1983), who states that experts are knowledgeable about their domain and they know how to apply and use the knowledge more effectively.

Hinds (1998), states that “experts are more thorough, more principled and more accurate in performing tasks within their domain of expertise” (p. 3). This is attributed to people taking a number of paths when evaluating information. These paths include pattern detection, use of principles, time spent on analysis, organisation of tasks and speed of task completion. Hinds (1998), infers that experts use these paths in a more effective manner than novices do. However experts are not always better suited to all tasks than novices. Hinds (1998) states that in tasks involving judgement and mathematical models, experts are generally not as accurate or even any better than novices carrying out the same tasks.

Buehler et al. (1994) in their study about how people predict time completion, hypothesised that

1) “People under-estimate their but not others completion time” (p. 366) and
2) “People focus on plan based scenarios rather than on relevant past experiences while generating their predictions” (p. 366)

The results from the study supported both these hypotheses. A number of reasons have come up as to why experts tend to underestimate. The major reason that has been documented is that as one becomes an expert they tend to break the view of the task down into smaller categories. Langer and Imber (1979) find that this “overlearning” has “debilitating effects” (p. 2014) for experts, as they found that as a task is repeated, each repetition leads closer and closer to a state of “mindlessness” (p. 2015). It becomes a process where the person knows how to complete the process but forgets about the steps that make up the performance of the process, hence leading to underestimation of the task they are carrying out. Another reason that has been documented is that people use a heuristic ofanchoring and adjustment (Nickerson et al., 1987). Hinds (1998) defines these “a process whereby people anchor on some value and then adjust that value to attain an estimate” (p. 3). It is the adjustment of the value that causes one to underestimate. This is found to be true when it comes to experts estimating novices. Experts estimate that novices lack the knowledge or experience of a particular task and under estimate what novices may be capable of. Expertise is something that usually occurs over time, according to Kahneman and Tversky (1979), and experts use
this experience over time to provide a best guess, estimate, or prediction relating to an uncertain aspect. Experts see novices as people without any experience hence, together with the above mentioned reasons expertise may lead to the overestimation of the time in which novices perform and complete a task (Hinds, 1998).

Expertise can be considered to be relative. An expert in one field or one specific aspect of a field may be novice or intermediate in other fields or other aspects of a field. Expertise can also vary among experts. In a research carried out by Kozma and Russell (1997), novices are defined as undergraduate chemistry students and professional chemists as experts.

Interestingly, Bereiter and Scardamalia, (1993) comment that that as the difficulty of a task increases, expert performance decreases as people choose to address the problems of their field at the upper limit of the complexity they can handle. They go on to say that “the career of the expert is one of progressively advancing on the problems constituting a field of work, whereas the career of a nonexpert is one of gradually constricting the field of work so that it more closely conforms to the routines the nonexpert is prepared to execute” (p. 11). However this experiment did not address this facet as the focus of the study was on simple everyday tasks that would not require one to be an expert to be able to perform the tasks effectively.

Santhanam and Wiedenbeck (1993) found that novices in the use of software encounter problems because they lack the knowledge about computers, their case in particular being word processing. This study links closely to the study that was carried out as the study also looked at word processing tasks. Two paradoxes are presented by Carroll and Rosson (1987), cited in Santhanam and Wiedenbeck (1993). These paradoxes are specifically for the learning of computers. The first paradox, the production paradox, claims that people’s strong desire to accomplish useful work is connected to their goals; hence people tend to learn about the computer and the appropriate software so that they can achieve their goal. However this does not mean they learn the most efficient way of completing the task. Also they will tend to focus only on the process needed to solve the task rather than learn the other properties of the program. The other paradox is that people naturally try to make sense of what they are learning by making reference to what they already know. This can have adverse effects especially when the user makes analogies that are not completely appropriate to the new situation. Santhanam and Wiedenbeck (1993) assert these paradoxes are general characteristics regardless of the level of expertise. The paradoxes deal with the learning of a process and the completion of a task with using analogies the production paradox will not be
observed in this study, however the second paradox will be simpler to observe and will be noted in this study.

With this in mind, in this research Comp101 participants are classified as novice users, to the particular environments they will be introduced to. We aim to see, by the use of questionnaires, there may be a correlation, positive or negative, between how people perceive their expertise and their actual performance of a task. The design of the experiment will be detailed later in the study.

2.2 Previous Experimental Methods

Gong and Salvendy (1995), in their study implemented one of the more promising approaches in carrying out the comparison between GUI and CLI's. They observed how the subjects learn the UNIX commands used to complete common file processing and directory manipulation tasks. The study was conducted over six consecutive practice sessions using four interface conditions: (1) menu only, (2) command line only, (3) user choice of menu or command line, and (4) automatic shifts from menu to command line, based on a user's improving performance with the menu interface. Recognition was programmed with the menu interface, and recall was programmed with the command line interface. Apparently, participants first sought the correct UNIX command to complete a task within an online help utility, and then attempted to input the correct command under the various interface conditions (Durham and Emurian, 1998).

In a study carried out by Kissel (2001) using participants with varying levels of computer experience, it was found that about 42% preferred the list box interface as the best, and about 38% preferred CLI’s, the difference was due to users not being able to complete the given tasks. As it can be seen from that study, there is not much of a difference between, CLI and list box. The low difference indicates that there was then still a number of people who prefer the use of CLI's. Hanson et al. (1994) found that a key reason for the demise of CLI’s was that users are able to make errors more readily because there is a lack of feedback. However a study carried out by Zanino et al (2001) showed that, contrary to popular belief, GUIs are not universally easy to use; certain types of individuals are likely to find them easier to use than others. Organisational roles and management initiatives can also influence perceptions of ease of use. The findings of this study were echoed by another study by Andre and Wickens (1990). They found that on comparing five system prototypes with a different interface that the interface with the worst preference ratings actually produced the best performances.
GUI's may not be better than CLI's, in all cases. Rohr and Kappel (1984) and Kacmar (1989), found there was no overall improvement in performance was noted between icon-based system over a text-based system. Nielsen (1987), when observing novice users using a Macintosh, noted the difficulties they had with the interface. On the other hand, Margano and Schneiderman (1987), when comparing MS-DOS commands with the Macintosh interface reported that the Macintosh interface led to faster task completion, fewer errors, and more subjective satisfaction. Similar results were observed by Svendsen (1993), whose study found that GUI's were found to be easier to learn and subjects preferred them to a CLI. The findings in both studies were attributed to user belief that

1) it is easier to learn and use commands and procedures on a GUI as it requires very little memorisation on the part of the users; and

2) the presence of the mouse increases the speed of performance and eliminates the errors associated with typing commands.

The findings also suggested that ease of use is enhanced through opportunities for self-training rather than traditional, formal training.

Whiteside et al. (1985) found that interface style was not an important predictor of performance or preference in their study. The interface style that was predicted as easiest for new users was actually the hardest and vice versa, GUI's were seen as being cumbersome and slow at that time, the processor speeds of the PC's available at that time were a major influential factor in how the subjects in this study reacted to the different interfaces. It was concluded that interface style, as manifested in real systems, is simply not an important factor for human factors of computer systems. They felt that interface styles do not, in themselves, solve old human factors problems (Whiteside et al, 1985). Hence the way an interface is designed is more important than the style of the interface chosen.

GUI's have been found to be effective on word processing tasks (Card et al., 1983; Jordan, 1992), file manipulation tasks (Morgan et al., 1991), and database retrieval tasks (Te’eni, 1990). These studies measured different combinations of factors, and they found that in most cases the GUI was only better in certain measures and not all. However, Eberts and Bittianta (1993) found that there were benefits in speed and understanding but not accuracy (1993), Morgan et al (1991) found benefits in accuracy and satisfaction but not speed while Frese et al. (1987) found that performance benefits are felt by regular users of an interface rather than beginners. An interesting note is that Erbets and Bittianta (1993) agree with the Hutchins et
First Impressions of Operating System Styles Affect Usability

P. Chianahal

al (1986) theory, saying that people prefer to use GUI's when the problem they are dealing
with is concrete and in spatial terms. However, if the problem is abstract then GUI's are not
normally preferred. Sekinger (1997) divided his subjects into three groups, namely expert,
average, and novice. He looked at two sections that an interface may influence, learning and
interface capabilities. He found that novice users ranked CLI's and GUI's higher in some
aspects than average and expert users. In others, expert users scored CLI's and GUI's higher
in some aspects than average and novice users. This is attributed to the fact that as people
progress from being novice to expert their expectations, feelings, and knowledge of both
types of interfaces change due to level of experience (Sekinger, 1997).

Many of the studies have pointed out that CLI's are good if one can specify exactly and
accurately what they want to do, because they can check what they have done easily and
because they can repeat complicated analyses either exactly or with simple changes of
parameters. However, Unwin and Hoffman (2001) give a variety of reasons that have lead to
the downfall of CLI's.

1) “Even experts mistype.” (p. 3)
2) “The commands are/may not be intuitive” (p. 3);
3) “They are suited for modelling and not for interactive graphics” (p. 3).

Such interfaces give the user the impression of working when in fact they are striving to
remember syntax and not concentrating on the task at hand (Unwin and Hoffman, 2001). This
process is described by Tognazzini (1992), as an abandonment of the primary task to give
attention to the secondary task. That is where GUI’s are found to be more advantageous, as
they are found to be intuitive, easy to learn, requiring recognition rather than recall (Unwin
and Hofmann, 2001). However, like CLI’s, GUI’s have their shortcomings. Unwin and
Hofmann (2001) maintain the two main ones being:

1) “They are not precise enough” (p. 3);
2) “They are difficult to design well” (p. 3).

Aprile (2004) states that, “GUI’s are great when you don’t really know what you’re doing
and are only going to do it once”. Kernighan (2002) suggests that the problem with GUI’s is
“finding the right buttons to push is like playing Adventure; the interfaces are just as
burdensome as CLI’s, save that one can in theory find the treasure by sufficient
exploration. With CLI’s, one needs a manual” (p. 23).
2.3 Summary

This review has shown the various findings with regards to comparing GUI's against CLI's. Studies have investigated user retention, recall, cognitive psychology, speed of processing and more. The studies are all aimed at finding which interface could be regarded as a good one. It would be safe to say that what is good depends on the needs and experience of the user, what they have to do, what facilities they have (hardware, other software, availability of help and advice) and on their style of working. The studies mentioned have given the advantages and disadvantages between GUI's and CLI's as they have observed them in different scenarios. Both interfaces can be improved, but the general consensus is that the better-looking interface does not always produce the best results. Whether one is an expert or a novice, there is always something that we as humans have not come across and the reaction to the new domain may affect their future interactions with it.

This study will differ from what has already been done. The studies that have been previously conducted have had experts being pitted against novices, which could be a reason why most of the studies show a tendency towards a particular interface style of another. In this study we aim to take a group of subjects that have relatively the same competency of computer use to allow for a more meaningful comparison. The similarity in the competencies will reduce the chances of level of expertise becoming an influential factor in the findings.

The previous studies have essentially been carried out when the GUI first came out, and since then the interface has evolved, as has the actual hardware. By using the more recently developed GUI's, to compare against the visually unchanged CLI, we will be able to establish if previous findings still hold today.

Hence the aim of this study is to use what has already been found, but bring it up to date using the modern type interfaces. We hope the findings will either support or refute the hypothesis
3. Methodology

The study is intended to observe if there is some influence of OS styles on the usability and perception of the use of computers. An interesting point to note is the perception of expertise against the actual performance of tasks. The study is a phenomenological study, a study in which we are attempting to understand people’s perceptions, perspectives and understandings of a particular situation (Leedy and Ormrod, 2001). By finding multiple perspectives of how users react to different operating system style, we can make some generalisations of how the styles affect reactions. This section will discuss the methodology of the study.

3.1 Sample

This study drew on novice users, which has been defined earlier as students currently enrolled in COMP101. The section about the difference between a novices and expert showed how the term novice can be defined with regard to the context it is being used in. The study was conducted over two days, during the normal running of COMP101 lab sessions in EAST CAL. Selection of students in the class was randomised and was essentially on a volunteer basis, no participants were forced to take part in the study. The study had to be performed over two days, so as to get a representative sample group. The optimum number was thought to be between 40 and 50 participants, i.e. about 10-12 participants per OS. Sample size is critical as it provides a basis for finding any trends that maybe present. The complete sample size used at the end of the data collection process numbered 60 individuals. This was made up of 15 individuals per OS. This sample size allows us to identify any trends, disparities and allow for retrieval of relevant information but insufficient to draw conclusive results.

3.2 Instrumentation

Initially it was proposed to derive information for six different terminals namely, Windows XP, Linux, DOS, UNIX, MacOS and Mac Terminal. However due to unforeseen circumstances I had to omit two terminal styles, i.e. Linux and UNIX. This would not hamper the study as it was regarded that the remaining four terminals would suffice in this study. A further change from the proposal was the use of a Windows 95 environment rather than a XP environment.

Two machines were used; a Digital laptop (model TS30G) running Windows 95 and MS-DOS version 6.22, and a Macintosh G3 running MAC OS X and Mac Terminal 1.3.1 (v82). Both machines were running with the same or similar types of software which would not affect data collection.
Time capturing software added to the machines. To ensure accuracy the time capturing software was correct down to the second. The tools used for the different OS styles were:

- Windows: A freeware spy software program, 007 Spy Software (version 3.4).
- Dos: a batch file implemented by the researcher.
- Mac OS X and Terminal: A freeware script modified to suit the requirements of this experiment, (REAL BASIC TimeTracker).

The study looked at both intra and inter interface comparisons. Figure 3 shows the aspects the study will be looking into. This is done so to ensure that before a comparing CLI's and GUI's we can group the DOS and Mac Terminal together generically as CLI's and be able to group Windows 95 and Mac OS X as GUI's.

![Diagram](image)

Figure 3 The areas the study aims to cover. Inter- PC vs. Mac and Intra GUI vs. CLI

### 3.3 Data Collection

Before participants were asked to carry out any of the tasks, they were given an information sheet detailing the basics of the study and a consent form to sign (Appendix A). A brief explanation was given and any questions were answered before the test started. No personal information about the participant was collected.
A task sheet was given to the participants (Appendix C). It contained a pre-experiment questionnaire (Appendix B), the tasks to be carried out and then by a post-experiment questionnaire (Appendix D). Each task sheet had a unique identifier that made no reference to the individual it was used to allow collection and later analysis of data to be made easier. Participants were assigned randomly to the different OS machines.

The aim of the pre-experiment questionnaire was to obtain the participants’ computer usage and the perception of their skill of use with regard to different operating system styles. Each participant in this study was asked to carry out simple tasks on the system they were assigned to. The study required subjects to carry out common tasks. Common tasks have been chosen as these are the tasks that most computer users will be familiar with. Previous studies (Buckelteiner and Estabrock, 1994) have tested users with software programs that are not intuitive, programs that people need to learn and then try to use. The learning process can play a big part in the findings. If the subjects are not interested in the software being used as the testing medium, then they are less likely to be as responsive, e.g. giving a subject an art program to try, but their interest lies in music. The tasks were the same across all the different interface styles.

The tasks involved creating a directory, (an interesting note is that most of the participants did not know what a directory was), open a text editor, perform simple text manipulation, save the file and exit the editor (See Appendix C for an example of the task sheets). The timers only started once the participants began to carry out the task. During the task participants were told to ask questions only if they were completely lost as this would affect the timing. On completion of the tasks the time capture was stopped and the participant completed the post-experiment questionnaire. The post-observation questionnaire obtained information on what the participant thought about the task they had just performed, as well as their perception about the interface and the time they took.

The text editors used in this study were the simple editors that come with the respective systems i.e. Notepad for Windows, SimpleText for Mac OS X, Editor for Dos and the Terminal itself for the Mac Terminal. These editors were chosen as the capabilities of each editor is very similar and each of the editors will not have a huge impact on the performance of the running of the systems.
Two participants performed the experiment at any one time, each one on a different interface i.e. one would be on the Mac Terminal and the other on Dos. No users would use more than one of the test interfaces. On completion of the experiment, data would be collected and the times were noted and logged on a separate PC, loaded with appropriate software required by the researcher.

3.4 Justification of questions

Apart from the tasks that were carried out by the participants data collection was supplemented in the form of answers to a questionnaire. As mentioned earlier, two questionnaires were used a pre questionnaire and a post-questionnaire. This section will give a brief justification of the questions.

3.4.1 Pre- Observation Questionnaire

1. Participant's Sex
   Male       Female

This question was asked as initially the researcher thought there may be a relationship between gender and the perception, however this was dropped as an analysis of the data collected it was found there were insufficient numbers of males and females to carry out a feasible comparison.

2. Do you normally use computers much?
   Yes       No

3. How often do you use one?
   Seldom     Weekly     Daily

4. Do you enjoy using computers?
   Yes       No

These questions aimed to obtain a basic background of the participant’s computer usage. This is because it is assumed that the answers to these questions may influence how participants carry out the tasks and will affect their view about the interface they use.
First Impressions of Operating System Styles Affect Usability

5. Have you ever used different computer interfaces?
   Yes  No

6. How would you rate yourself as a user of these Operating Systems? (1 = total beginner 5 = expert)
   MS Windows  1  2  3  4  5
   DOS          1  2  3  4  5
   MacOS        1  2  3  4  5
   Linux        1  2  3  4  5

Usage of different interfaces and the perception of the participant’s view of their expertise in each of the interfaces show how participants regard themselves as users of systems. The fact that subjects may have used different interfaces could be indicative that they should be somewhat quicker at tasks. This perception could then be tested to see if there is a correlation between perceived expertise and the way the tasks are carried out.

7. Do you know the difference between a Graphical User Interface and a Command Line Interface?
   Yes  No

As this study is comparing GUI’s and CLI’s this question aims to observe if people do know the difference between the two, or are these two words just terms that are known in the computer community. An interesting comparison could be made using the answers from question 6 to see if there is any correlation between people who regard themselves as relatively good users of both GUI’s and CLI’s and see if they do know the difference between the two.

3.4.2 Post-Observation Questionnaire

1. Which Operating System did you use?
   MS Windows  Dos  MacOS  MacConsole

This question showed whether the participants really knew what OS they were using, initially it was thought that it may have some influence on the findings but it was seen to not provide sufficient information to make a meaningful observation.
2. Did you enjoy using this Interface?
   Yes  No

The satisfaction, which is hard to quantify, is addressed by this question. Even though the tasks were very quick we can use this to identify any relationship (if any) between this and the time taken to do the tasks. As mentioned earlier it is difficult to measure feelings, this questions aims to obtain the participants initial perception of the interface.

3. How long do you think the task took you to do?
   >10mins 10mins 5mins <5mins Not Sure

This question attempts to address the time perception taken on the tasks. The answers given here will be compared to the actual time recorded. This will give an interesting measure if what was mentioned in the literature review, that normally people under-estimate the time they take to complete a task.

4. Were the tasks simple to do and understand?
   Yes  No

This question aims to show that the structure and wording involved with the tasks did not affect the outcome of the experiments, i.e. not affecting satisfaction, speed or usability.

5. Do you think that you were more accurate using this OS
   Yes  No

6. Do you think you were more efficient (faster) than using some other OS?
   Yes  No

These two questions address the users association and perception of other OS styles with regard to the one they had just used. It will enable us to draw some inferences as to what participants thought of the style they had just used, especially if they hadn’t used this type of interface before.

7. Has this experiment shown you things you didn’t know about other interface system?
   Yes  No

When using something new the assumption is that one will learn something about it, be it bad or good. This question will also show if users did learn anything or not, this could tie in with their level of perceived expertise.
3.5 Data Analysis

SPSS statistical software was used for all quantitative analysis. Independent t tests and One way ANOVA was carried on various combinations of the questions. Only the relevant statistical data was calculated. Statistical visualisations were also plotted to illustrate any observable trends. The full details of the results and analysis will be covered in the following section. Qualitative analysis was a bit more difficult to measure as there is no easy way to quantify just data, general statements were made about these findings that’s encompass the findings which may relate to reviewed literature. There is no correct method of dealing with any missing data, inputting missing data will influence the outcome of the results, and will affect the overall conclusion. Only complete relevant questions were used when conducting analysis. Careful consideration was implemented to ensure all the questions were answered, so as to reduce the chances of experimental error that could arise due to missing data.
4. Results

The data collected from the study was used to analyse two areas. Across the different operating system styles i.e. GUI and CLI as well as between the interfaces. The primary areas of interest are Time Taken to complete the tasks, Perceived Expertise, Perceived Time Taken and Satisfaction. These are the focus points of the study and this section will quantify and analyse them.

4.1 Numerical Findings

Before doing the statistical analysis, a simple summary was done showing some of the numeric findings that of the data collected. Table 3 shows these findings.

<table>
<thead>
<tr>
<th>Table 3 Average time, longest time, shortest time and the overall time taken on the four in interfaces</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Average Time Taken (secs)</strong></td>
</tr>
<tr>
<td>Dos</td>
</tr>
<tr>
<td>Longest Time (secs)</td>
</tr>
<tr>
<td>Shortest Time (secs)</td>
</tr>
<tr>
<td>Sum of Total Time (secs)</td>
</tr>
</tbody>
</table>

It can be seen that overall task time on GUI’s took much longer to complete than the CLI’s. The average time of the CLI’s compared to the GUI’s further emphasises this difference with a numerical difference of approximately 31.20 seconds (using the average of the averages, 209.42 vs. 178.22). The slowest task completion time was with the Windows interface (486.60 secs) and the quickest task completion took place on the Dos interface (86.40 secs). The raw data seem to imply that CLI’s fare better with regards to time.

An important point must be mentioned before further analysis is conducted, one of the questions asked was “Were the tasks simple to do and understand?”, the answer to this question is important as it may have an effect on the results, the researcher was hoping that most participants would find that the task instructions were indeed simple. It was found that 91.7% of the participants found the tasks simple to follow and understand the further 8.3% may have found the tasks harder to understand where the participants came from backgrounds in which English is not the first language, hence they may have found some difficulty with understanding the tasks. With a high percentage agreeing that the tasks were straightforward,
it can safely be concluded that the actual task and the instructions should have no bearing on the results.

4.2 Statistical Findings
An independent t-test was performed using the user interfaces as the independent variable and the dependent variables used for this evaluation was the time taken as this would be an indicator if there is a difference between the two interfaces. The results from the t-test for the CLI's showed that even though the mean time is greater for Dos than the Terminal the difference between the two is not significant (t = 0.132, n.s.), this allows us to group Dos and Terminal as CLI’s. The same was done for the GUI's and it was found that though the mean time of the Mac OS X was found to be greater than Windows, the difference between them was not significant (t = -0.123, n.s.), like the Dos and Terminal. This finding allows us to group Windows and Mac OS X as GUI’s. These findings show that there is minimal statistical difference to prevent us from grouping these interfaces. It also allows the subsequent statistics to occur across the two styles of interface. As stated in the Methodology section this is important to do as it allows to state that it is valid to group all the CLI or GUI results together.

Before a comparison of the different interface styles can be done, it is important to ensure that there is no significant difference between similar interfaces. This is to ensure that both Dos and Terminal can be grouped as CLI’s in this experiment and Windows and Mac OSX and be referred to as being GUI’s.

4.2.1 Time Taken
Graph 1 illustrates the findings mentioned Table 3, showing that the CLI’s have a mean time taken that is less than the GUI’s. However a more accurate picture can be given by taking the log of the times and trying to see if there was really any difference statistically between the times of the different interfaces and the total time taken. Using Levene’s test of Equality for Error Variances it was found that the relationship between interface types on time taken has a significant value of 0.62. Although this indicates some significance, the value is regarded by statisticians as one that can not be taken to indicate any meaningful relationship.

Graph 1. Mean total time (sec) against the type of interface used
A one way ANOVA was run to ensure the findings for Levene's test were correct. It was found that $F(3,59)=1.139, n.s.$ supporting the fact that in this study there is no significant relationship between the type of environment and the time taken.

The distribution for the 4 interfaces (Figure 3) show that with the GUI environments most of the participants complete the tasks relatively fast, however those that take a longer time it cause the distribution to skew, hence the higher mean times. Of the CLI's, Dos shows a normal distribution, which indicates that the task completion time on average of all the participants was evenly spread, with a lot more participants completing the task around the mean time. The terminal times are more random than any of the other interfaces thus no clear distribution can be identified.

*Figure 4. Histograms illustrating the distribution of all the interfaces with regards to time.*
Graph 2. Time taken on the two interface styles

Graph 2 illustrates the comparison of the times of all the participants between the two interface styles, it can be seen that users of the CLI were generally much quicker in completing the tasks irrespective of which CLI was used (DOS or Terminal). Even though the statistical run shows there is no statistical relationship between the time taken and the interface style used, it can be clearly seen that users of the CLI interface were generally much faster, hence suggesting that there must be some relationship though it may not be linked to time. There are three exceptions above 300 seconds, this might support the notion that GUI’s can take longer, but not supported statistically with this sample size.

4.2.2 Time Perception

One of the questions asked participants to indicate how long they felt they took to complete the task. It is interesting to note that when comparing the type of interface used, the majority of participants felt they took less than 5mins, irrespective of the interface used (Table 4). Statistically there appears to be no relationship between the type of interface one used and the time that was perceived to have been taken (t = 1.304, n.s.).

From Table 5, below it can be see that the participants whose perception times are the highest and lowest on the different OS styles show that with in the different ranges i.e. <10mins, 10mins, 5mins and >5mins, there is a huge variability in the times, this might be due to perceived expertise, (which will be covered later in the chapter).
Table 4. Number of participants and the time they thought they took to complete the task

<table>
<thead>
<tr>
<th>Time taken</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;10mins</td>
<td>2</td>
</tr>
<tr>
<td>10mins</td>
<td>5</td>
</tr>
<tr>
<td>5mins</td>
<td>17</td>
</tr>
<tr>
<td>&lt;5mins</td>
<td>33</td>
</tr>
<tr>
<td>not sure</td>
<td>3</td>
</tr>
</tbody>
</table>

An interesting point to note are that on the Mac Os ratings most people rated themselves in the 5min and < 5min range, this is notable as on the other interfaces the predictions are more varied across the spectrum. This table does not indicate the numbers in each group according to the interface.

Earlier it was said that there is no relationship between interface type and perceived time, to ensure that there could be no other cases where a relationship could be found a One way ANOVA was carried out, it was found that F(3,59) = 2.240, n.s.), supporting the previous finding that there is no relationship with regards to perceived time and interface style.

Graph 3 illustrates this fact, as it can be seen clearly that irrespective of whether the interface was a GUI or a CLI participants still gave almost similar answers to the time perception. Graph 4 below shows there is a negative correlation with regards to the log of the total time.

Table 5. Maximum and minimum times within each type of interface grouped by the answers given on the post questionnaire.

<table>
<thead>
<tr>
<th>Windows</th>
<th>&lt;5mins Max</th>
<th>Min</th>
<th>5mins Max</th>
<th>Min</th>
<th>&gt;5mins Max</th>
<th>Min</th>
<th>Not Sure Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>TotalTime</td>
<td>486.60</td>
<td>372.00</td>
<td>240.00</td>
<td>156.00</td>
<td>305.40</td>
<td>121.00</td>
<td>196.60</td>
<td>122.40</td>
</tr>
<tr>
<td>Seconds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacOs</td>
<td>TotalTime</td>
<td>364.20</td>
<td>176.20</td>
<td>262.20</td>
<td>142.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seconds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dos</td>
<td>TotalTime</td>
<td>197.40</td>
<td>122.20</td>
<td>199.80</td>
<td>86.40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seconds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terminal</td>
<td>TotalTime</td>
<td>236.20</td>
<td>121.20</td>
<td>306.00</td>
<td>120.60</td>
<td>124.80</td>
<td>124.80</td>
<td></td>
</tr>
<tr>
<td>Seconds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
actually taken and perceived time, across the different interfaces \((r = -0.191)\). This is quite unusual as the earlier table indicates that participants were relatively close to time perception vs. actual time taken. There could be some confounding factors, which will be looked at in the next chapter.

**Graph 3** Showing the perception of time taken for participants on GUI and CLI

**Graph 4.** The correlation between the logs of actual time (seconds) against perceived time taken
4.2.3 Perceived expertise

A number of statistical analyses can be conducted with regards to perceived expertise, the one that is most interesting to us is perceived expertise against the type of interface used. I will also look at perceived expertise against time taken and perceived time taken. The perceived expertise as can be expected is lower for CLI’s than for GUI’s.

The questionnaire included Linux as it is another OS that is based on GUI, this was to see if people who may regard themselves as good users of Linux may have a different opinion of other GUI’s, however, no statistical analysis of any relevance was carried out on those responses as the Linux OS had been hardly used my most of the participants, only 5 participants had indicated they had used a Linux OS.

![Perceived Expertise on the Two Interface Styles](image)

**Graph 5** A participants perceived expertise of use on a GUI and a CLI

Graph 5 shows that numerically a majority of user’s perceptions of expertise was lower on the CLI’s while a higher number felt they could be regarded as “better” experts on the GUI interface style. It was found that the mean expertise was higher for GUI’s than on CLI’s. This was expected as most participants have been exposed to the GUI interface more than a CLI type interface. However the interesting point is that as this experiment was aimed to target novice users, the mean average (Table 6) found can clearly show that most of the participants can be considered as novices, with few exceptions. There is no relationship between the type of interface used and the participant’s perception of their expertise, (GUI F(3,59) = 1.63, n.s., CLI F(3,42) = 1.09, n.s.). An indication that the participants can be regarded as novices can also be observed as was stated earlier by the fact that most participants did not know what a directory was, a fact that someone taking the COMP101 course would be required to know.
Table 6. Mean Expertise of GUI's and CLI's

<table>
<thead>
<tr>
<th></th>
<th>GUI</th>
<th>CLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.85</td>
<td>1.63</td>
</tr>
</tbody>
</table>

Perceived expertise with relation to the time taken shows that across the two styles there may be a relationship between perceived expertise and time taken. There is a negative correlation that exists between the expertise and the time taken to complete the task (r = -0.218, n.s). However, the relationship, in this study, is not significant because it only suggests that as expertise increases the times taken to complete the task would decrease. An interesting indication from the statistics is that most people who regarded themselves as experts on the GUI took relatively longer than those who ranked themselves lower. The overall difference in expertise between GUI's and CLI's was found to be 2.70 as compared to 1.50, not a very large difference. However, it is interesting to note that people working on the GUI's found it much harder to predict their time, irrespective of their expertise, with their perceptions lying across the different ranges whereas the participants using the CLI environments seem to have been able to perceive their time within a narrower range.

Table 7. Means of the expertise grouped by perceived time taken between GUI's and CLI's

<table>
<thead>
<tr>
<th>How long do you think the task took?</th>
<th>GUI</th>
<th>CLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;10mins</td>
<td>2.50</td>
<td></td>
</tr>
<tr>
<td>10mins</td>
<td>3.00</td>
<td>1.50</td>
</tr>
<tr>
<td>5mins</td>
<td>2.40</td>
<td>1.25</td>
</tr>
<tr>
<td>&lt;5mins</td>
<td>2.75</td>
<td>1.86</td>
</tr>
<tr>
<td>Not Sure</td>
<td>3.00</td>
<td></td>
</tr>
</tbody>
</table>

4.2.4 Satisfaction

Satisfaction is hard to quantify. An attempt will be made to establish if there is any relationship between the style of interface and satisfaction. As this is a very subjective variable to gauge, the analysis for satisfaction will be conducted on answers from questions 2, 5 and 6 of the post questionnaire. This is because satisfaction is defined as “The human experience of being filled and enriched by their experience.” (DeVinney, 1997), with a study of this scale, and with the questions asked, the three questions in will give the best indication.
Table 8. Number of participants and their responses to the questions on enjoyment, accuracy and efficiency for GUI's and CLI's

<table>
<thead>
<tr>
<th></th>
<th>GUI</th>
<th>CLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyment</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>12</td>
</tr>
<tr>
<td>Accuracy</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>Efficiency</td>
<td>16</td>
<td>14</td>
</tr>
</tbody>
</table>

Raw data (Table 8) shows that in general participants felt that GUI’s were to be more enjoyable and accurate. With the CLI’s the raw data suggests a very narrow difference of opinion, a difference that may not be statistically significant. The efficiency findings show clearly that CLI participants felt that were not as efficient using the CLI environment, claiming they were more competent on some other type of interface style. However this is not as evident with regards to those using the GUI styled interfaces, the difference only being about 6%, however it must be noted that the GUI styles were Mac OS X and Windows and on further analysis it is found that generally most participants did not enjoy (9), find themselves accurate (13) or think there were efficient (9) using the Mac Os comparatively participants using the Windows OS only 4 participants said they did not enjoy using the interface, no participants indicated they were not accurate with windows and 5 said they felt they were not efficient using the Windows interface.

Graphs 6 & 7: The percentage of participants and how they found the different interface styles with regards to enjoyment, accuracy and efficiency, all very subjective measures.
A univariate suggests that there is a significant relationship between the type of interface style and the enjoyment that participants felt. Univariate tests were also carried out on accuracy and on efficiency and it was found for accuracy $F(3,59) = 4.43$, n.s. and for efficiency $F(3,59) = 1.83$, n.s. the p-value for the accuracy (0.07) was only just approaching significance it would not be suitable to suggest that there may be no relationship. Hence a t-test was conducted to see if there could be any significance in the relationship. Carrying out independent t-tests on style against accuracy on the two styles, it was found there is no significance. ($t = -0.273$, n.s.).

To find if perceived expertise has any influence on how participants felt about the interface styles, a cross tabulation was carried out to describe the association between the subjective feelings and the supposed expertise. With enjoyment it was established that across the two interface styles, taking perceived expertise ratings of 4 and 5 to be high, it was found that with GUI's 39 people felt they were on the high range of expertise. Of the 39, 58.9% felt they had not enjoyed utilising the GUI they had used. In comparison only 10 people felt they could regard themselves as relative experts on the CLI style, and of these 10, 60% enjoyed the CLI style. With regards to the lower end of expertise, the findings showed that 44.6% of the users of the GUI had not enjoyed the interface style while 55% of CLI participants responded in the negative with regards to enjoyment. No significant relationship was found between perceived expertise and enjoyment.

With accuracy it was found that users of the GUI's, irrespective of expertise, did not rate themselves as being accurate on the interface they used. But, of the 6 people that regarded themselves a 5 on expertise on a GUI, all 6 felt they were also accurate elsewhere. Despite the raw data suggestions, expertise and accuracy was found statistically to have a positive correlation with interface style (0.125), indicating that there may be some relationship. A One way ANOVA showed that despite the finding that there maybe some correlation between accuracy and perceived expertise, there is no significant relationship between how one views their accuracy with regards to how they perceived their expertise.

Efficiency saw most users with the higher perception of expertise claim that they were not as efficient using the interface styles, whilst those on the lower end of the expertise scale alleged the opposite. 73.2% of the people who ranked themselves as low expertise agreed that they did find the interface style they had used as being more efficient, of these 39.4% used a CLI's and 33.7% were participants on a GUI's.
The illustration (Graphs 8 & 9) shows that there is an apparent trend. It seems that enjoyment, accuracy and efficiency may be linked; the graphs show that generally those with a perceived expertise of 3 agreed they enjoyed the interface style as well as its accuracy. However, it can be seen for those who said no, a lot of them felt they had expertise of 4. The graphs do not show which interface styles were measured, it shows the number of participants on all the interface styles.

Participating in the study, however, we found that even though this was a very simple task, the majority of people felt they learnt something, only 31.6% of the participants felt they had learnt nothing. For completeness, a statistical run showed that (F(3,59)=3.23, p<0.05) there is a strong relationship between the interface style and the fact that participants felt they had learnt something using the interface styles.

Before undertaking the task, one of the questions asked if participants knew what the difference is between a GUI and CLI. None of the participants asked for a clarification of what the question was asking, indicating that they knew was being asked of them. The findings showed that 76.7% of the participants did not know the difference between a GUI and a CLI whilst the rest said they did (23.3%). The interesting point to note is that 71.7% of the participants said they had used a CLI (DOS), calculated using the answers to question 7 of the pre-questionnaire. The fact that even though the majority claimed that they did not know the difference between the two interface styles, yet had used the two interface styles is an interesting point for discussion.
4.3 Summary

The results show that users (slightly) prefer a GUI interface style, but are more productive on the CLI. The results also confirm the fact that the participants in the study can be regarded as novices.
5. Discussion

The hypothesis of this study is that providing novice users with certain operating system styles (GUI or CLI) will influence their perception of computer usability. Novice users are expected prefer and be more productive using a GUI interface, the results give interesting readings; this section will aim to clarify the results.

The results reveal a number of key points; it is seen that initially CLI’s are quicker performance wise. This can be attributed to the fact that there was no recall required. According to Galitz (1993) and Unwin and Hofmann (2001) the downfall for CLI’s has been the fact that they require recall of commands in order to carry out a task. This study eliminated the need for recall as the commands were clearly laid for participants in the task sheet. Since the tasks were the same for the GUI and CLI, it suggests that if users of CLI’s are given the commands for the tasks rather than the need to recall them they would be able to perform tasks quicker than on GUI’s. The times may have also been slower on the GUI’s as CLI processes are less processor intensive. According to Afinogenov (2003) even though today’s hardware largely renders this point arguable. CLI’s incur much less size and speed overhead than GUI’s. It was attempted to eliminate the effect of processor speeds from being a factor, by using computers of the same processor specifications although, it must be stated it may have played some part, e.g. when users created directories in this study, compared to the one line required on the CLI’s the GUI’s need a couple of steps to complete this simple task.

Perception of time taken on each of the tasks is linked to what was found in the literature, especially with regards to novice users. Earlier it was stated that Buehler et al. (1994) found that experts in particular “under-estimate their own but not others completion time” (p. 366). The study found that most of the novice participants estimated their times very close to the actual times they had taken. Buehler et al. (1994) carried out their study on experts, and it is interesting to point out that in our findings we have shown that novices unlike experts don’t seem to underestimate time taken, rather they seem to have a better perception of time. However the negative correlation observed in this sample group (Graph 4) indicates that the relationship of actual time taken (seconds) and perceived time means that if people were to take longer in completing a task their perception of the time taken would decrease. This finding does not fit with any of the literature and logically does not make sense. According to Niven (2004), it is impossible to formulate precise recommendations based on these
correlation results as the sample size was not large enough to allow for any serious biases to be noticed.

Expertise was considered in this study even though the study was focused on novice users. This is because even though people may be regarded as novices, within the novice group there will be a range of perceptions as to how individuals would rate themselves. From the results it was seen that even though there were a number of people who regarded themselves as relative experts on a GUI, the mean expertise across the GUI style was 2.85. This is an indication that most of the users regarded themselves as relative novices on both GUI and on CLI. Perception of one expertise can be used to show that even though people may regard themselves as different levels of expertise, their actual performance may not indicate their perceptions e.g. in our study one of the participants rated him/herself as an expert (5) on the Windows operating style but their time taken to complete a task was almost 4 minutes, which is the same time as those that regarded themselves of lesser expertise. But there are some results which indicate the findings are not very conclusive if perceived expertise could be a good indicator of how one would perform on an interface style. It must be remembered that one’s perception of expertise is made up of a number of factors some of which may not be related to the actual task at hand but rather to an individuals own self judgement levels. Referring to Table 7 it can be seen that the mean evaluations follow no set trend, with a (relatively) high proportion corresponding to a moderately long completion time. This would tend to support a hypothesis like: “A given task may confirm novice status irrespective of perceived expertise.” This would follow the experience had by those teaching novice students; students may overestimate their ability yes will not be able to complete the work as they don’t really know what they are doing.

The results showed that there was a conflict between the answers participants gave for whether they knew the difference between a GUI and a CLI and their use of the different interface styles. The results show that even though most of the participants had used a CLI (71.7%) and GUI (100%) a large number of participants (76.7%) said they did not know the difference between the two interface styles. This indicates that users may have used the different interface styles, without understanding the inherent differences of the styles. Surprisingly even though the look and feel of the two styles is very different, and should have indicated that there is some difference between the styles, it seems that the participants did not warrant that as a good enough reason to say they were different. However it could also be argued that the participants may regard the differences as being on the technical front and not
just on the aesthetics and design. Amusingly, after one of the participants had completed the study, they commented that even though they were told the difference between the two interface styles in one of their COMP 101 classes, they could not recall them.

The fact that participants claimed that they did not know the difference between the two interface styles would not have affected the overall results, but serves to reinforce novice status. Also this is an interesting finding and would be interesting to observe what participants would interpret by the statement “knowing the difference between” the interface styles.

Usability of any system is linked to the satisfaction a user feels after using a system. There are a number of heuristic measures proposed by various researchers in the field (Nielsen and Mack, 1994 and Tognazzini, 1992). This study addressed usability by looking at enjoyment, accuracy and efficiency. These factors were compiled after study of the heuristic lists made by the above mentioned authors. From the results it can be seen clearly that there is a relationship between the interface style and the participant’s enjoyment. GUI’s are found to be more enjoyable to use than CLI’s. There are a number of factors that may be the reason why there is a slightly higher preference a GUI styled interface. A GUI is seen to be more aesthetically pleasing than CLI interface, a user is able to relate to pictures better than they are to words. The saying “A picture is worth a thousand words” is proven to be correct in this example, because a GUI uses concepts that are familiar to people while CLI’s use a language based on computer terminology which is difficult for most users to relate to. The metaphor basis of building GUI’s mean that a user is presented with objects on the screen that they are able to decipher using real world experiences, e.g. the folder system in Windows replicates what would happen in real life when one is filing documents.

Another factor that affects enjoyment is Engagement. Hutchins et al (1986) and Frolich (1997) both agree that users prefer using systems that give the user a feeling of control over the system. Another point raised in the literature is the effect of mental effort on a users’ perception of a system, the higher the mental toil the less a user would enjoy using the system (Norman, 1986). GUI’s seem to “engage” users of them much better than CLI’s and the mental toil of using a GUI because of the use of metaphors, is much less than on CLI’s.

These are factors that lead to a GUI being perceived as more enjoyable to use than a CLI. Apart from enjoyment, accuracy and efficiency were also considered, both factors which lead to satisfaction and ultimately to usability. As another perception related question, the aim was
to detect any preferences towards a particular interface style as being accurate and efficient in the eyes of the participant. From the results it can be seen that with regards to accuracy the GUI interface style is perceived to be more accurate than a CLI. However the difference between the two is interface styles is very slight, and there is no overwhelming majority that prefer one style as being more accurate over the other. This indicates that in this group approximately 50% of users of the GUI and CLI test interfaces found themselves to be more accurate on the respective interface style, than on any other style they had used. Some individuals perceive themselves as being more accurate than they actually are, it a question again of how people scale themselves with regards to what they regard as being accurate.

With efficiency it is clearly seen that participants using the CLI did not find themselves as being efficient on that interface style. Efficiency is one of the requirements on Nielsen’s (1994) list of a system with good usability. With the GUI interfaces, the difference in the number of participants who perceived themselves as being more efficient using this interface style is very slight i.e. 46% No and 56% Yes. Like accuracy it is very hard to make a convincing statement with the results obtained, although it seems that the GUI styled interface would be regarded as being better in both cases.

According to Svendsen (1991), when a system is termed as being user friendly it means that users of the system are able to learn the functionality of the interface fairly quickly and are able to use this functionality to solve problems. However the task in this study was not long enough for participants to be able to learn the full functionality of the systems, but Svendsen (1991) goes on to state that the assumption that a user friendly interface makes for better problem solving over a less user friendly interface, and this indicates that usability somehow aides in problem solving. Our study shows that this is not the case as, although the GUI style was perceived to be more user friendly, task completion times suggest that the CLI style is better for performance. If we had tested problem solving we might hypothesise the same results would be seen.

It was found from the study that a majority of the users of the Macintosh GUI found it not user friendly. 60% stated they had not enjoyed the experience and felt they were not efficient using it. An astounding 86% felt they were not accurate on this interface. This is surprising as 13 of the 15 people (86%) that carried out the task on the Macintosh GUI had used a Mac before, hence one would assume that the previous use would give them some background knowledge on how to use a Mac GUI and not give this result. A factor that may have resulted
in the findings obtained with regards to the Mac is that despite previous use of a Mac, it is assumed that the Windows environment is seen as a benchmark with which most people today view other OS styles and may play a role in participant perceptions.

From the results of this study, the preference ranking of the operating styles is
1. Microsoft Windows,
2. Dos,
3. Mac Terminal and
4. Mac OS
The fact that the Microsoft products are preferred over the Apple product, may be coincidence. The more remarkable fact is that participants seemed to prefer both the CLI styles over the MAC OS X. This is remarkable as the aim of the Apple is to be as user friendly interface. The fact that Microsoft Windows is seen as the preferred interface style can be attributed to the Windows interface being the most popular and widely used interface by a majority of computer users. Even if people have not used a windows system before, they will have encountered it in some manner or another. As at September 2004 almost 90% of OS’s in use today are based on the Windows family, with only 2.6% of OS’s being from the Macintosh family. (W3Schools, 2004).
6. Conclusion

The hypothesis for this study was that providing novice users with certain operating system styles (GUI, CLI) will influence their perception of computer usability. Novice users are expected prefer and be more productive using a GUI interface. The study provided some support for the conjecture that novice users would be prefer using the GUI styled interface, however their productivity measured as the time taken to perform a task, was not found to conclusively show that the productivity of novice users was significantly better on a GUI platform. Rather it was found that productivity was better on the CLI, as discussed in section 5.

Revisiting the hypotheses that
1. Novice users will prefer a GUI to a CLI and
2. Novice users will be more productive using a GUI than a CLI

We find that users (slightly) prefer a GUI interface style, but are more productive on the CLI. Hence hypothesis one has been confirmed, but hypothesis two has not been confirmed. Rather novice subjects were more productive using the CLI interface.

The premise of the hypothesis was about perception of usability. The findings show that perception of usability is strongly related to the type of interface style one uses. Usability itself was not measured, however aspects that would make up usability such as accuracy, efficiency, learning and enjoyment were measured, and using those as indicators, it is found that there was an influence on the users perceptions of usability. The general impression is that GUI’s are easier to use than CLI’s. Even though the participants are novices, the GUI style is one that most people are aware of, and advertising campaigns of most software packages sell the idea that a better GUI makes for better usability. However there are a number of confounding factors that may have influenced the findings of this study.

6.1 Confounding factors

The results showed that in this study most of the analyses did not show many significant relationships. This may be correct for the study, however to make a more accurate comment that would be appropriate for a broad-spectrum of the computer user population we need to mention some of the confounding factors that may have influenced the findings of this study.
• **Sample Size**

Earlier in this report it was stated that for this study 40-50 participants would be the optimum, 60 participants were used in total. However as this group was divided into 4 categories, it meant only 15 people were tested on each interface type. This number is not large enough to allow one to observe any noticeable trends. According to Niven (2004), a study of this nature would require a minimum of 50 people to be tested on each interface. This would enable one to observe any trends clearly. It would also allow for statistical analysis to show any significant relationships.

• **Questions**

The questions are a confounding factor, as most of the questions had Yes or No answers. This may not be a good thing as sometimes things are not a clear cut Yes or No. A rating scheme would have been better. However, the problem with a rating scheme is that perception across the range may differ e.g. for a rating scheme of 1 2 3 4 5, the perceived gap between 1 and 2 may be different when compared to the perceived gap between 4 and 5 (Niven, 2004). On balance a rating scheme would have been better than a YES or No answer scheme.

Another point is that more questions could have been asked about how respondents felt about the interface they had just used, and maybe some open ended question about what they felt was good or bad about the interface style they had just used. This may give some insight as to what people find appealing (or otherwise) with the two interface styles.

• **Tasks**

Only one task was carried out. If more tasks were carried out it would have taken away any biases towards any one program. Word processing was tested here and the timings maybe affected by participants typing speeds. More tasks across other areas, it would give a better picture of a user’s performance with the interface. One short task is not a true representation of how users may feel about a particular interface as it is not enough time for them to get to grips with the interface style.

• **Number of Variables**

The study had too many variables (expertise, learning, enjoyment, four different interfaces). This can be attributed to the fact inexperience on the researchers’ part. Due to the number of variables, each could have played some part in the overall results. Each variable would have to be studied independently to see if they could affect the usability of the interface.
• Subjectivity

Subjectivity is defined as one that is “open to different interpretations based on prior experience or expertise. Having no definitive or specific interpretation. Having a broad range of answers depending on the basis for the response.” (Technical Forecasts Limited, 2004). The nature of this study, as with most HCI studies, is very subjective. Perceptions of the participants are influenced by many factors, and in a subjective study, the influence of these factors may affect the results. Each individual perceives things in a unique manner. With a larger sample size we could negate some of the factors. According to Meloche (2000), using the Q methodology one can reveal subjective views that individuals hold on particular topics. In future studies this technique may be used.

6.2 Future work

Future work could look at how frequent and expert users would respond to the same scenario. Future work could also look at more specific interface types, and compare them, rather than using a wide range, e.g. Focusing on MS DOS vs. Windows. Further research could also look at a comparison of errors between the interface styles. One of the guiding questions in this study was if the criteria used to measure reactions be accurate. As it can be seen the criteria are accurate to some extent, however, there is still lots of room for improvement and future work could look at trying to find more accurate methods of measuring subjective criteria.

To conclude, the results from this study serve to support what has been found before and suggests that there is no direct relationship between preference and performance. However to test the usability of a program, many other factors should be looked at, as they could all be contributing factors determining the usability of an interface style. Addressing the confounding factors in future research may provide more conclusive evidence affecting the design of future interface styles. People like attractive looking screens but seem to work better on screens with a minimalist design.
References


“First Impressions of Operating-System Styles Affect Usability”

P. Chhanabhai


Niven, B. (2004). Interview. [Interview with Brian Niven, 10 September, 2004].


First Impressions of Operating System Styles Affect Usability

P. Chhanabhai


Appendices
Appendix A

8th June 2004

IMPRESSIONS OF OPERATING SYSTEM STYLES AFFECT USABILITY
INFORMATION SHEET FOR PARTICIPANTS

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate. If you decide to participate we thank you. If you decide not to take part there will be no disadvantage to you of any kind and we thank you for considering our request.

What is the Aim of the Project?
This project is being undertaken as part of the requirements for the Postgraduate Diploma in Science.
The major aim of this project is to see whether the style of operating system you use determines your ability to perform common computing tasks.

What Type of Participants are being sought?
Participants will be made up of COMP101 students. COMP101 students may be considered novice users.

What will Participants be Asked to Do?
Should you agree to take part in this project, you will be asked to:
Perform some simple tasks on the computer, and to give some feedback about your thoughts on the tasks completed. The tasks will involve text editing and file management. These tasks should not take more than 10 minutes of your time.
Please be aware that you may decide not to take part in the project without any disadvantage to yourself of any kind.

Can Participants Change their Mind and Withdraw from the Project?
You may withdraw from participation in the project at any time and without any disadvantage to yourself of any kind.

What Data or Information will be Collected and What Use will be Made of it?
Data will be collected by using a closed questionnaire. There will be a pre-test question part, as well as a post-test question section. The information collected will be about prior computing knowledge (the pre-questionnaire), followed by the impressions of the tasks carried out (the post-questionnaire).
The questionnaire will have a unique ID on it for the purpose of linking the questionnaire to the results of the tasks that will be carried out. No information that could identify a participant is collected.
Results of this project may be published but any data included will in no way identify any specific participant.
You are most welcome to request a copy of the results of the project should you wish.
The Information collected will be used for the support (or otherwise) of the researchers hypothesis. Access to the data will be made to researchers, supervisors and co-ordinators. The data collected will be securely stored in such a way that only these people will be able to gain access to it. At the end of the project any personal information will be destroyed except that, as required by the University's research policy, any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed.

What if Participants have any Questions?
If you have any questions about our project, either now or in the future, please feel free to contact either:-

Prajesh Chhanabhai
Department of Information Science
University Telephone Number: N/A

or

Samuel Moyle
Department of Information Science
University Telephone Number: 4798198
IMPRESSIONS OF OPERATING SYSTEM STYLES AFFECT USABILITY

CONSENT FORM FOR PARTICIPANTS

I have read the Information Sheet concerning this project and understand what it is about. My questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I know that:-
1. My participation in the project is entirely voluntary;
2. I am free to withdraw from the project at any time without any disadvantage;
3. The data will be destroyed at the conclusion of the project but any raw data on which the results of the project depend will be retained in secure storage for five years, after which it will be destroyed;
4. The results of the project may be published and that every attempt will be made to preserve my anonymity.

I agree to take part in this project.

................................................................. .................................................................
(Signature of participant) (Date)
Appendix B

Pre-Observation Questionnaire
(Please Circle your Selection)

1. Participant Sex
   Male    Female

2. Do you normally use computers much?
   Yes    No

3. How often do you use one?
   Seldom  Weekly  Daily

4. Do you enjoy using computers?
   Yes    No

5. Have you ever used different computer Operating Systems?
   Yes    No

8. How would you rate yourself as a user of any Operating Systems you have used? (1 = total beginner 5 = expert)
   MS Windows  1  2  3  4  5
   DOS          1  2  3  4  5
   MacOS        1  2  3  4  5
   Linux        1  2  3  4  5

9. Do you know the difference between a Graphical User Interface and a Command Line Interface?
   Yes    No
Appendix C

Tasks

Windows

The operating system that you are using is Windows 95, a Graphical User Interface (GUI). You are asked to:

- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

1.

- Locate the folder named 480 which can be found on the desk top
- Open the sub folder named Windows
- Create a Folder in there and call it by the number that is on the top of this sheet
- Close the Window

2.

- Open NotePad (the icon can be found on the desk top)
- Create a text Document with the following text:

  “This is some text. Now I have to save it. My test number is xxx”

  (where xxx is the number at the top of this sheet.)

- Save your work in the Folder you just created. Name your file <xxx.txt>.
- Close Notepad
MacOs

The operating system that you are using is the Macintosh Operating System. It is a Graphical User Interface (GUI). You are asked to:

- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

Before starting the tasks

❖ Press the Start Time button that you see on the Timer Application.
❖ Click anywhere on the desktop

1.

➢ At the top of the Screen you will see a Finder Tool Bar
➢ Click File, New Folder
➢ A folder will pop up on the desktop. Name it the number that is on the top of this sheet.

2.

➢ Open SimpleText (it can be found on the desktop)
➢ Create a text Document with the following text.

“This is some text. Now I have to save it. My test number is xxx”

(where xxx is the number at the top of this sheet.)
➢ Save your work in the Folder you just created. Name your file xxx.txt.
➢ Close SimpleText

Before ending: Press the Stop Time button on the Timer Application
Tasks

**DOS**
The operating system that you are using is Microsoft Disk Operating System (DOS). It is a Command Line Interface (CLI). You are asked to:
- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

1. The screen you will have in front of you will have a command line looking like this
   \texttt{C:~}\texttt{480\Dos}\texttt{~}
   - To Create a directory, type \texttt{md} \texttt{space} followed by the name of the directory, in this case it will be the number at the top of the sheet.

2. With \texttt{C:~}\texttt{480\Dos}\texttt{~} being displayed, we will now create a file
   - Type \texttt{TIMER}
   - The Dos Editor should open (A Blue Screen)
   - Create a Document with the following text (just start typing)
   "This is some text. Now I have to save it. My test number is xxx"

   (where xxx is the number at the top of this sheet.)

   - Save your work by naming it the xxxx.txt in the Folder you just created.**

   **The mouse does not work in dos mode, so the Save will have to be done using Keyboard Commands.
   - To Save press ALT+F
   - go to Save As
   - Press ALT+D and go to the Directory you created earlier,
   - Press ALT+N and type in the file name (the number at the top of the sheet).
   - Press Tab three times, till the OK button is selected, then press enter.
   - The file is now saved.

   - To quit, press ALT+F to get the file menu and choose the <Exit> option. Press <enter>, and you will be exited from the program.
**Tasks**

**MacOs – Console**

The operating system that you are using is the Macintosh Console Operating System. It is a Command Line Interface (CLI). You are asked to:

- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

Before starting the tasks

- **Press the Start Time button that you see on the Timer Application.**

2. The screen you will have in front of you will have a command line looking like this:
   ```
   [ou075217:~/480]mshlp$
   ```

   - To Create a directory, type `mkdir <space>` followed by the name of the directory, in this case it will be the number at the top of the sheet.

2. With `[ou075217:~/480]mshlp$` being displayed, we will now create a file

   - Type `touch` followed by the number that is on the top of the sheet
   - Type `echo` followed by:

     "This is some text. Now I have to save it. My test number is xxx"

     (Where xxx is the number at the top of this sheet.)

     The file is saved.

Before ending:

- **Press the Stop Time button on the Timer Application**
Appendix D

Post-Observation Questionnaire

(Please Circle your Selection)

5. Which Operating System did you use?
   MS Windows   MS Dos   MacOs   Mac Console

6. Did you enjoy using this Interface?
   Yes         No

7. How long do you think the task took you do?
   >10mins  10mins  5mins  <5mins  Not Sure

8. Were the tasks simple to do and understand?
   Yes         No

5. Do you think that you were more accurate using this OS style
   Yes         No

6. Do you think you were more efficient (faster) than using some other OS style?
   Yes         No

8. Has this experiment shown you things you didn’t know about other Operating Systems?
   Yes         No
“After climbing a great hill, one only finds that there are many more hills to climb. I have taken a moment here to rest, to steal a view of the glorious vista that surrounds me, to look back on the distance I have come. But I can rest only for a moment, for with freedom comes responsibilities, and I dare not linger, for my long walk is not yet ended.” - Nelson Mandela
PART B

RESEARCH ARTICLE
First Impressions of Operating System Styles Affect Usability

Prajesh Chhanabhai

University of Otago

Prajesh Chhanabhai is studying for a Masters in Science in Information Science with an interest in the usability of interfaces, and the affect of computers on human thinking.
ABSTRACT

The patterns of behaviour that people develop to work successfully with complex information technology are likely themselves to be complex. The beginning of interface style followed this train of thought in developing the Command Line Interface, complex to develop and complex in its use. However, in recent years information system interface design has become increasingly dominated by the use of Graphical User Interfaces, with the majority of systems relying on a Microsoft Windows based structure. This study attempts to find out how the two different interface styles would affect how novice users use them when given a word processing task. The study was conducted using two interfaces based on a graphical style and two command line type operating systems. The study examined how quickly the participants performed the task on the different interface styles, and used questionnaires to supplement the quantitative findings. Preference versus performance was studied and the findings are in sync with what other researchers have found. The results indicate that better usability may not mean better performance.
CONTENTS

1. INTRODUCTION
   1.1. General Background
   1.2. Hypothesis
   1.3. Variables
       1.3.1. Independent Variables
       1.3.2. Dependent Variables
2. REVIEW OF LITERATURE
   2.1. Novice vs. Expert
   2.2. Previous Studies
3. METHOD
4. RESULTS
5. DISCUSSION
6. CONCLUSION
1. INTRODUCTION

So many operating systems and so many graphical desktop environments... System developers face many challenges and decisions as they analyze and create designs for the various types of interfaces around. One of the biggest factors for them to consider is how the system will function and operate for the intended user (Merchant, 2002). Jefferson (2002) postulates that the two main reasons for continued “lousy interfaces” are the reduction in the importance of end-user experience and designing based on the requirement of operating systems rather than those of the end-user. Manes (2000) suggests that there is a growing public sentiment of interface dissatisfaction as most of the interface styles anger and confuse users.

There have been numerous studies carried out over the years, comparing Command Line Interface (CLI) against the Graphical User Interface (GUI). This study looks into some of the research that has occurred in this field and extracts the essential findings from these studies. This study attempts to support previous findings.

1.1. General Background

According to Smith (1998), billions of dollars have been invested into the Information Technology (IT) sector. Studies have shown that white collar productivity is still low (Smith, 1998). The relationship between IT investment and productivity is alarming to organisations and individuals alike. Smith (1998) concludes that considerable time and money is being spent to acquire, train and use such technology. The key barrier to productivity is that many users of IT do not actually know how to use it and many people are scared of using such technology. Some contributing factors are the inherent usability of the system (Schneiderman, 1982), functionality it provides (Davies, 1989) and the man-made interface (Schneiderman, 1982). It can be argued that of the three factors the one that affects IT users most directly is the man-made interface as its design determines eventual user acceptance and utilization (Davis and Bostrom, 1993). Graphical user interfaces (GUIs) are the most pervasive interface at least partly because of conventional wisdom about their ease of use. Such an assumption may have been kindled by vendor claims about the inherent usability of such interfaces. Research on the productivity gains from GUIs has yielded mixed results (Zantos, Prasad, Agarwal, 2001).

The purpose of this study is to assess the usability of four interfaces that are used today. The type of interface may be linked to productivity, and this study looks into the relationship between performance and preference as these may infer a link to productivity.

1.2 Hypothesis

It is hypothesized that providing novice users with certain operating system styles (GUI and CLI) will influence their perception of computer usability. Novice users are
First Impressions of OS Styles Affect Usability

expected to prefer and be more productive using a GUI interface. The hypothesis can be broken into two sections.

1. Novice Users will prefer a GUI to a CLI.

2. Novice users will be more productive using a GUI than a CLI.

1. 3 Variables

The variables investigated in this study were

1.3.1 Independent Variables

This will be an active independent variable as the researcher will be able to monitor its effect on the dependent variables. The independent variables in this study are the Operating Systems or Environments. This can be further classified as four independent variables GUI preference, GUI performance, CLI preference and CLI performance.

1.3.2 Dependent Variables

The key dependent variables studied was the usability, impressions and quality of human computer interaction. These were obtained by looking at perceived time to perform a task, actual time to perform a task, perceived expertise on the interface style used. Essentially we are indicating ease of use, speed and perceptions.

2. REVIEW OF LITERATURE

The usability of menu interfaces compared to command line interfaces began to appear in literature after the introduction of direct manipulation input devices, such as the mouse (English, Engelbart and Berman, 1967). The purpose of these direct manipulation devices was to simplify human-computer interactions (Perry and Voelcker, 1989; Shneiderman, 1983). It was considered that menu interfaces required a loss of a cognitive effort than that of CLI’s. However despite that assumption, Liu (1997) and Shneiderman (1998) comment that novice computer users are seen to prefer using a menu driven interface whilst expert users prefer to use the power of the command line. That conclusion is based upon the assumption that recognising a correct command is easier than recalling a correct command. Paap and Roske-Hofstrand (1988) acknowledge that the usability advantage of one interface over another is influenced by many interacting factors, including task complexity, user objectives, user facility in manipulating different input devices, and user knowledge of a particular application domain.

According to Rautenberg (1996) the classification of interfaces is important; two quantitative measures that are suitable to allow this classification are functional feedback and interactive directness. Using these it is possible to classify the most common interface types: batch, command, menu and desktop. The command line interface is characterised by high interactive directness, but this interface type has a very low amount of visual feedback. Only GUIs are found to support the user with sufficient visual
First Impressions of OS Styles Affect Usability

feedback and with high interactive directness. To make this classification as understandable as possible, Rauterburg (1996) describes the three classified interfaces with an abstract schema (see Figure 1). “FIGURE 1 ABOUT HERE”

One important method of assessing how good interfaces are is to measure their usability. Nielsen and Levy (1994) suggest that measuring usability falls into two broad categories:

1) subjective user preferences which measures how much the users like the system.

2) objective performance measures which measures how capable the users are at using the system.

One would expect that there would be a direct positive correlation between subjective preference and objective performance, as people can be expected to prefer doing their job well and to use computers with interfaces that help them rather than hinder them (Nielsen and Levy, 1994). This is true for most cases; however there are examples when users do not like using a system as it slows down their ability to do a task, even though they like its overall layout. For example MacLean, Barnard and Wilson (1985), found that subjects preferred the slower of two data entry methods as long as it was not 20% slower than the faster method.

Despite the view among interface designers that recognition using a GUI is easier than recall using a command line (Galitz, 1993), this view is found to be the exception rather than the norm (Greene, Gould, Boies, Rasarnay and Melson 1992). Svendson (1991) declares that the reason people associate a user-friendly interface as one that makes it a good problem-solving tool is because people assume that:

1) using an interface and solving problems compete for the same cognitive resources; and

2) a user-friendly system somehow facilitates problem solving.

According to Svendson (1991), both assumptions are hard to measure and are not warrantable as many factors affect the user friendliness of a program which will affect problem solving skills as well.

2.1 Novice vs. Expert

When a novice ventures into a new domain it can be observed that the user makes many errors and requires a lot more time to solve problems. However over the course of time and through continuous use the user becomes faster and more accurate. This transition from a novice to a user that is able to perform functions quicker and with accuracy maybe said to be a transition from being a novice to becoming an expert user. According to Kolodner (1983), experts are knowledgeable about their domain and they know how to apply and use their knowledge more effectively.
First Impressions of OS Styles Affect Usability

Baehl, Griffin and Ross (1994), in their study about how people predict time completion, hypothesised that

1) “People under-estimate their but not others completion time” (p. 366) and

2) “People focus on plan based scenarios rather than on relevant past experiences while generating their predictions” (p. 366)

The major reason that has been documented is that as one becomes an expert they tend to break the view of the task down into smaller categories. Experts estimate that novices lack the knowledge or experience of a particular task and under estimate what novices may be capable of. Expertise is something that usually accrues over time, according to Kahneman and Tversky (1979), and experts use this experience over time to provide a best guess, estimate, or prediction relating to an uncertain aspect. Experts see novices as people without any experience hence, they overestimate the time in which novices would perform and complete a task (Hinds, 1998).

Santhanam and Wiedenbeck (1993) found that novices in the use of software encounter problems because they lack knowledge about computers, their case in particular being word processing. This study links closely to the study that was carried out as the study also looked at word processing tasks. Two paradoxes presented by Carroll and Rosson (1987) cited in Santhanam and Wiedenbeck (1993). These paradoxes are specifically for the learning of computers. The first paradox, the production paradox, claims that people’s strong desire to accomplish useful work is connected to their goals; hence people tend to learn about the computer and the appropriate software so that they can achieve their goal. However this does not mean they learn the most efficient way of completing the task. Instead they will tend to focus only on the process needed to solve the task rather than learn the other properties of the program. The other paradox is that people naturally try to make sense of what they are learning by making reference to what they already know. This can have adverse effects especially when the user makes analogies that are not completely appropriate to the new situation. Santhanam and Wiedenbeck (1993) assert these paradoxes are general characteristics regardless of the level of expertise.

2.2 Previous Studies

In a study carried out by Kissel (2001) using participants with varying levels of computer experience, it was found that about 42% preferred the list box interface as the best, and about 38% preferred CLI’s, the remaining 20% were unable to complete the given tasks. As it can be seen from that study, there is not much of a difference between CLI and list box. The low difference indicates that there was then still a number of people who prefer the use of CLI’s. Hanson, Kraut and Farber (1994) found that a key reason for the demise of CLI’s was that users are able to make errors more readily because there is a lack of feedback. However a study carried out by Zarine, Prasad and Agarwal (2001) showed that, contrary to popular belief, GUIs are not universally easy to use; certain types of individuals are likely to find them easier to use than others. Organisational roles and management initiatives can also influence perceptions of ease of use. The findings of
First Impressions of OS Styles Affect Usability

this study were echoed by another study by Andre and Wickens (1990). They found that on comparing five system prototypes with a different interface that the interface with the worst preference ratings actually produced the best performances.

GUI’s may not be better than CLI’s in all cases. Rohr and Keppel (1984) found no overall improvement in performance between icon-based system over a text-based system. Nielsen (1987), when observing novice users using a Macintosh, noted the difficulties they had with the interface. On the other hand, Margato and Schneiderman (1987), when comparing MS-DOS commands with the Macintosh interface reported that the Macintosh interface led to faster task completion, fewer errors, and more subjective satisfaction. Similar results were observed by Svendsen (1993), whose study found that GUI’s were found to be easier to learn and subjects preferred them to a CLI.

Whiteside, Jones, Levy and Wixon (1985) found that interface style was not an important predictor of performance or preference in their study. The interface style that was predicted as easiest for new users was actually the hardest and vice versa. GUI’s were seen as being cumbersome and slow at that time and processor speeds of the PC’s available at that time were a major influential factor in how the subjects in this study reacted to the different interfaces. They felt that interface styles do not, in themselves, solve old human factors problems (Whiteside, Jones, Levy and Wixon, 1985). Hence the way an interface is designed is more important than the style of the interface chosen.

3. METHOD

The study is based on the perceptions gained from novice users, in that regard, novice users in this experiment were obtained from the COMP101 course that is run at the University of Otago. COMP101 is an introductory course into computing. The course is aimed for students that find using computers and information technology challenging, lack exposure to use the essential computer applications, use computers unproductively and struggle to achieve goals, and more importantly do not understand what they are doing when using a computer.

The study was conducted over two days, during the normal running of COMP101 lab sessions. Participants were students in the class and it was on a randomized basis, essentially on a volunteer basis, with no participants forced to take part in the study. The sample size used for data collection process numbered 60 individuals. This was made up of 15 individuals per OS style. This sample size allows us to identify any trends, disparities and allow for retrieval of relevant information. Two machines were used: a Digital laptop (model TS300) running Windows 95 and MS-DOS version 6.22, and a Macintosh G3 running MAC OS X and Mac Terminal 1.3.1 (v82). Both machines were running with the same or similar types of applications which would not affect data collection.

Additional Software added to the machines included time capturing tools. Time was measured in seconds. The tools used for the different OS styles were: Windows: A freeware spy software program, 007 Spy Software (version 3.4). Dos: a batch file implemented by the researcher. Mac OS X and Terminal: A freeware script found on the
internet but modified to suit the requirements of this experiment, (REAL BASIC
TimTracker).

Pre-experiment and post-experiment questionnaires were used to collect data from the
participants. The aim of the pre-Experiment questionnaire was to obtain the participant’s
computer usage and the perception of their skill of use with regard to different operating
system styles. Each participant then carried out simple tasks on the system they were
assigned to. Common tasks have been chosen as these are the tasks that most computer
users will be familiar with. Previous studies (Buckleitner and Estabrook, 1994) have
tested users with software programs that are not intuitive, programs that people need to
learn and then try to use. The learning process can play a big part in the findings. If the
subjects are not interested in the software being used as the testing medium, then they are
less likely to be as responsive, e.g. giving a subject an art program to try, but their interest
lies in music (Hanson, Kraut and Farber, 1994). The tasks were the same across all the
different interface styles.

The tasks involved creating a directory, opening a text editor, performing simple text
manipulation, saving the file and exiting the editor (See Appendix for an example of the
task sheets). The timers only started once the participants began to carry out the task. On
completion of the tasks the time capture was stopped and the participant completed the
post-experiment questionnaire. The post-experiment questionnaire obtained information
on what the participant thought about the task they had just performed, as well as their
perception about the interface and the time they took. The text editors used in this study
were: Notepad for Windows, SimpleText for Mac OS X, Editor for Dos and the Terminal
itself for the Mac Terminal. Results were analysed using SPSS statistical software.

4. RESULTS

A simple summary was done to show some of the numeric findings of the data
collected. “FIGURE 2 ABOUT HERE”

From Figure 2 it can be seen that overall task time on GUI’s took much longer to
complete than the CLI’s. The average time of the CLI’s compared to the GUI’s further
emphasises this difference with a numerical difference of approximately 31.20 seconds
(using the average of the averages, 209.42 - 178.22). The slowest task completion time
was with the Windows interface (486.60 secs) and the quickest task completion took
place on the Dos interface (86.40 secs). The raw data seem to imply that CLI’s fair better
with regards to time. An important point must be mentioned, one of the questions asked
was “Were the tasks simple to do and understand?” the answer to this question is
important as it may have an effect on the results. It was found that 91.7 % of the
participants found the tasks simple to follow and understand the further 8.3% may have
found the tasks harder to understand as a majority of the participants came from
backgrounds in which English is not the first language. With a high percentage agreeing
that the tasks were straightforward, we can safely conclude that the actual task and the
instructions should have no bearing on the results.
First Impressions of OS Styles Affect Usability

Time Taken

"FIGURE 4 ABOUT HERE"

Figure 4 illustrates the findings mentioned in Figure 2, showing that the CLI's have an overall mean time taken that is less than the GUI's. However a more accurate picture can be given by taking the log of the times and trying to see if there was really any difference statistically between the times of the different interfaces and the total time taken. Using Levene’s test of Equality for Error Variances it was found that the relationship between interface types on time taken has a significance value of 0.62. Although this indicates some significance, the value is regarded by statisticians as one that can not be taken to indicate any meaningful relationship. A One Way ANOVA was run to ensure the findings for Levene’s test were correct. It was found that F(3,59)=1.139, n.s. supporting the fact that in this study there is no significant relationship between the type of environment and the time taken.

"FIGURE 5 ABOUT HERE"

Figure 5 illustrates the comparison of the times of all the participants on the two interface styles, it can be seen clearly from here that users of the CLI were generally much quicker in completing the tasks irrespective of which CLI was used (Dos or Terminal). Even though the statistical run shows there is no statistical relationship between the time taken and the interface style used, it can be clearly seen that users of the CLI interface were generally much faster, thus for a more conclusive finding another test should be run with more participants.

Time Perception

One of the questions asked participants to indicate how long they felt they took to complete the task. It is interesting to note that when comparing the type of interface one used, the majority of participants felt they took less than 5mins, irrespective of the interface used. Statistically there appears to be no relationship between the type of interface one used and the time that was perceived to have been taken (t = 1.304, n.s.). An interesting point to note are that on the Mac Os ratings most people rated themselves in the 5min and < 5min range. This is notable as on the other interfaces the predictions are more varied across the spectrum.

A One Way ANOVA was carried out, it was found that F(3,59) = 2.240, n.s.), supporting the previous finding that there is no relationship with regards to perceived time and interface style.

Perceived expertise

The perceived expertise, is lower for CLI’s than for GUI’s. The questionnaire included Linux. This was to see if people who may regard themselves as good users of Linux may have a different opinion of other GUI’s. However, no statistical analysis of any relevance was carried out on those responses as the Linux OS had been hardly used by most of the participants.
First Impressions of OS Styles Affect Usability

“FIGURE 6 ABOUT HERE”

Figure 6 shows that numerically a majority of user’s perception of expertise was lower on the CLI’s while a higher number felt they could be regarded as “better” experts on the GUI interface style. We found that the mean expertise was higher for GUI’s than on CLI’s, this is something that was expected as most participants have been exposed to the GUI interface more than a CLI type interface. There is no relationship between the type of interface used and the participant’s perception of their expertise, (GUI F(3,59) = 1.63, n.s., CLI F(3,42) = 1.09, n.s.). There is a negative correlation that exists between the expertise and the time taken to complete the task (r = -0.218, n.s.). The overall difference in expertise between GUI’s and CLI’s was found to be 2.70 as compared to 1.50, not a very large difference.

Satisfaction

Satisfaction is hard to quantify because this is a very subjective variable to gauge. We will try to establish if there is any relationship between the style of interface and satisfaction. Satisfaction will be measured using the answers from questions 2, 5 and 6 of the post questionnaire.

Raw data (see Figure 3) shows that in general participants felt that GUI’s were to be more enjoyable and accurate. Between the CLI’s the raw data suggests a very narrow difference of opinion, a difference that may not be statistically significant. The efficiency findings show clearly that CLI participants felt that were not as efficient using the CLI environment, claiming they were more competent on some other type of interface style.

“FIGURE 3 ABOUT HERE”

However this is not as evident with regards to those using the GUI styled interfaces, the difference only being about 6%, however it must be remembered that the GUI styles were Mac OS X and Windows and on further analysis it is found that generally most participants did not enjoy (9), find them selves accurate (13) or think there were efficient (9) using the Mac Os.

A univariate analysis shows that though it is hard to find a clear relationship using raw data, statistically we find that F(3,59) = 2.24, p<0.05, suggesting there is a significant relationship between the type of interface style and the enjoyment that participants felt.

Univariate tests carried out for accuracy and efficiency found that, for accuracy F(3,59) = 4.43, n.s. and for efficiency F(3,59) = 1.83, n.s. The p-value for the accuracy (0.07) was only just approaching significance it would not be suitable to suggest that there may be no relationship. Hence a t-test was conducted to see if there could be any significance in the relationship between accuracy and interface style. This showed there was no significance in the relationship (t = -0.273, n.s.).
5. DISCUSSION

The results reveal a number of key points. With the result that CLI's are quicker performance wise. According to Galfre (1993) and Unwin and Hofmann (2001) the downfall for CLI's has been the fact that they require recall of commands in order to carry out a task. This study eliminated the need for recall, across both the CLI's as the commands were clearly laid for them in the task sheet. Since the tasks were the same for the GUI and CLI, it suggests that if users of CLI's are given the commands for the tasks rather than the need to recall them they would be able to perform tasks quicker than on GUI's. The times may have also been slower on the GUI's as CLI processes are less intensive when it comes to using the processor. According to Affroganov (2003) even though today's hardware largely renders this point arguable, CLI's incur much less size and speed overhead than GUI's. We had tried to eliminate the effect of processor speeds being a factor, although, in our discussion we must state it may have played some part, e.g. when users created directories in this study, compared to the one line required on the CLI's the GUI's need a couple of steps to complete this simple task.

Perception of time taken on each of the tasks is linked to what was found in the literature, especially with regards to novice users. Buehler, Griffin and Ross (1994) found that experts "under-estimate their own but not others completion time" (p. 366). From the study we found that most of the novice participants estimated their times very close to the actual times they had taken. Buehler, Griffin and Ross (1994), carried out their study on experts and it is interesting to point out our findings show that novices, unlike experts, don’t seem to underestimate time taken. Rather they seem to have a better perception of time.

From the results it was seen that even though there were a number of people who regarded themselves as relative experts on a GUI, the mean expertise across the GUI style is 2.85. This is an indication that most of the users were relative novices on both GUI and on CLI. Perception of expertise can be used to show that even though people may regard themselves as different levels of expertise their performance may not indicate their perceptions. For example in our study one participant rated themself as an expert on the Windows operating style but their time taken to complete a task was the same time as those that regarded them selves of lesser expertise. Some users whose perception of expertise was on the lower end tied in with the actual time they had taken to do a task. Thus it is not conclusive if perceived expertise could be a good indicator of how one would perform on an interface style. It must be remembered that ones perception of expertise is made up of a number of factors some of which may not be related to the actual task at hand but rather to an individuals own self judgement levels.

From the results it can be seen clearly that there is a relationship between the interface style and the participant's enjoyment. GUI's are found to be more enjoyable to use than CLI's. There are a number of factors influencing the leaning towards a GUI styled interface. A GUI is perceived as being more aesthetically pleasing than a CLI interface, a user is able to relate to pictures better than they are to words. The saying "A picture is worth a thousand words" is proven to be correct in this example. This is because GUI's use concepts that are familiar to human beings whilst CLI's use a text
First Impressions of OS Styles Affect Usability

based on computer terminology which is difficult for most users to relate to. The metaphor basis of building GUI’s mean that a user is presented with objects on the screen that they are able to decipher using real world experiences, e.g. the folder system in Windows replicates what would happen in real life when one is filing documents.

From the results of this study, we found the preference ranking of the operating styles to be Windows, Dos, Mac Terminal and Mac OS. This is remarkable as the aim of the Apple is to be a user friendly interface. The fact that Microsoft Windows is seen as the preferred interface style, can be attributed to the Windows interface being the most popular and widely used interface by a majority of computer users, and even if people have not used a windows system before they will have encountered it in some manner or another. As at September 2004 almost 90% of OS’s in use today are based on the Windows family with only 2.6% of OS’s being from the Macintosh family. (W3Schools, 2004).

6. CONCLUSION

The hypothesis for this study was that providing novice users with certain operating system styles (GUI, CLI) will influence their perception of computer usability. Novice users will prefer and be more productive using a GUI interface. The study provided some support for the conjecture that novice users would be prefer using the GUI styled interface. Productivity (time taken to perform a task) was not found to conclusively show that the productivity of novice users was significantly better on a GUI platform, rather we found that productivity was slightly better on the CLI. We found that perception of usability is strongly related to the type of interface style one uses. Usability itself was not measured, however aspects that would make up usability such as accuracy, efficiency, learning and enjoyment were measured. Using these indicators we find that there was an influence on the users perceptions of usability.

Future work could look at more specific interface types and compare them, rather than using a wide range, e.g. Focusing on MS DOS vs. MS Windows. Further research could also look at a comparison of errors between the interface styles. One of the guiding questions in this study was if the criteria used to measure reactions be accurate. As it can be seen the criteria are accurate to some extent, however, there is still lots of room for improvement and future work could look at trying to find more accurate methods of measuring subjective criteria.
NOTES

Acknowledgments. The author would like to thank Samuel Moyle for his supervision and knowledge imparted, Brendon Sly of TSG for the machines given to the author to allow testing, Colin Aldridge for his guidance through the research process.

Authors' Present Addresses. Prajesh N Chhanabhai, University of Otago, 60 Clyde Street, Dunedin, New Zealand, Email: chhpkr38@student.otago.ac.nz.
REFERENCES


First Impressions of OS Styles Affect Usability


First Impressions of OS Styles Affect Usability


APPENDICES

A1

Pre-Observation Questionnaire

(Please Circle your Selection)

1. Participant Sex
   Male
   Female

2. Do you normally use computers much?
   Yes
   No

3. How often do you use one?
   Seldom
   Weekly
   Daily

4. Do you enjoy using computers?
   Yes
   No

5. Have you ever used different computer Operating Systems?
   Yes
   No

6. How would you Rate yourself as a user of any Operating Systems you have used?
   (1 = total beginner 5 = expert)
   MS Windows
   1
   2
   3
   4
   5
   DOS
   MacOS
   Linux
   1
   2
   3
   4
   5

7. Do you know the difference between a Graphical User Interface and a Command Line Interface?
   Yes
   No
Windows

The operating system that you are using is Windows 95, a Graphical User Interface (GUI). You are asked to:

- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

1. 
   - Locate the folder named 480 which can be found on the desk top
   - Open the sub folder named Windows
   - Create a Folder in there and call it by the number that is on the top of this sheet
   - Close the Window

2. 
   - Open NotePad (the icon can be found on the desk top)
   - Create a text Document with the following text:
     “This is some text. Now I have to save it. My test number is xxx”
     (where xxx is the number at the top of this sheet.)
   - Save your work in the Folder you just created. Name your file <xxx.txt>.
   - Close Notepad
MacOs

The operating system that you are using is the Macintosh Operating System. It is a Graphical User Interface (GUI). You are asked to:

- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

Before starting the tasks

- Press the Start Time button that you see on the Timer Application.
- Click anywhere on the desktop

1.
- At the top of the Screen you will see a Finder Tool Bar
- Click File, New Folder
- A folder will pop up on the desktop. Name it the number that is on the top of this sheet.

2.
- Open SimpleText (it can be found on the desktop)
- Create a text Document with the following text.

"This is some text. Now I have to save it. My test number is xxx"
(where xxx is the number at the top of this sheet.)
- Save your work in the Folder you just created. Name your file xxx.txt.
- Close SimpleText

Before ending:
Press the Stop Time button on the Timer Application
First Impressions of OS Styles Affect Usability

A4

Tasks

DOS

The operating system that you are using is Microsoft Disk Operating System (DOS). It is a Command Line Interface (CLI). You are asked to:

- Create a Directory
- Make a file
- Do some simple text editing
- Save your Work.

1. The screen you will have in front of you will have a command line looking like this:
   C: \480 \DOS>>

   ➢ To Create a directory, type md space followed by the name of the directory, in this case it will be the number at the top of the sheet.

2. With C: \480 \DOS>> being displayed, we will now create a file
   ➢ Type TIMER
   ➢ The Dos Editor should open (A Blue Screen)
   ➢ Create a Document with the following text (just start typing)

   "This is some text. Now I have to save it. My test number is xxx"
   (where xxx is the number at the top of this sheet.)

   ➢ Save your work by naming it the xxxx.txt in the Folder you just created.**

   ** The mouse does not work in dos mode, so the Save will have to be done using Keyboard Commands.
   ~To Save press ALT+F
   ~go to Save As
   ~Press ALT+D and go to the Directory you created earlier.
   ~Press ALT+N and type in the file name (the number at the top of the sheet).
   ~Press Tab three times, till the OK button is selected, then press enter.
   ~The file is now saved.

   ➢ To quit, press ALT+F to get the file menu and choose the <Exit> option. Press <enter>, and you will be exited from the program.
First Impressions of OS Styles Affect Usability

MacOs – Console

The operating system that you are using is the Macintosh Console Operating System. It is a Command Line Interface (CLI). You are asked to:

- Create a Directory
- Make a File
- Do some simple text editing
- Save your Work.

Tasks

Before starting the tasks

◆ Press the Start Time button that you see on the Timer Application.

2. The screen you will have in front of you will have a command line looking like this
   {ou075217:~/.480]marrp%
   ➢ To Create a directory, type mkdir <space> followed by the name of the directory, in this case it will be the number at the top of the sheet.

2.
   ➢ With {ou075217:~/.480]marrp% being displayed, we will now create a file
   ➢ Type touch followed by the number that is on the top of the sheet
   ➢ Type echo followed by:

   “This is some text. Now I have to save it. My test number is xxx”

   (Where xxx is the number at the top of this sheet.)

   The file is saved.

Before ending:

◆ Press the Stop Time button on the Timer Application
First Impressions of OS Styles Affect Usability

Post-Observation Questionnaire
(Please Circle your Selection)

1. Which Operating System did you use?
   MS Windows    MS Dos    MacOs    Mac Console

2. Did you enjoy using this Interface?
   Yes    No

3. How long do you think the task took you do?
   >10mins   10mins  5mins  <5mins  Not Sure

4. Were the tasks simple to do and understand?
   Yes    No

5. Do you think that you were more accurate using this OS style?
   Yes    No

6. Do you think you were more efficient (faster) than using some other OS style?
   Yes    No

7. Has this experiment shown you things you didn’t know about other Operating Systems?
   Yes    No
First Impressions of OS Styles Affect Usability

FIGURE CAPTIONS

Figure 1. A classification schema of most common user interfaces. (Rauterburg, 1996.)

Figure 2. Average time, longest time, shortest time and the overall time taken on the four in interfaces

Figure 3. Number of participants and their responses to the questions on enjoyment, accuracy and efficiency for GUI's and CLI's

Figure 4. Mean total time (secs) against the type of interface used

Figure 5. Time taken on the two interface styles

Figure 6. A participants perceived expertise of use on a GUI and a CLI
First Impressions of OS Styles Affect Usability

FIGURES

Figure 1

[visual] feedback (FB)

<table>
<thead>
<tr>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>batch</td>
</tr>
<tr>
<td></td>
<td>menu interface</td>
</tr>
<tr>
<td></td>
<td>MI</td>
</tr>
<tr>
<td>high</td>
<td>command language</td>
</tr>
<tr>
<td></td>
<td>direct manipulation</td>
</tr>
<tr>
<td></td>
<td>DI</td>
</tr>
</tbody>
</table>
First Impressions of OS Styles Affect Usability

**Figure 2**

<table>
<thead>
<tr>
<th>(secs)</th>
<th>Dos</th>
<th>Terminal</th>
<th>Mac OS X</th>
<th>Windows</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Time</td>
<td>177</td>
<td>174</td>
<td>211</td>
<td>207</td>
</tr>
<tr>
<td>Taken</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longest Time</td>
<td>263</td>
<td>306</td>
<td>364</td>
<td>486</td>
</tr>
<tr>
<td>Shortest Time</td>
<td>86</td>
<td>115</td>
<td>142</td>
<td>121</td>
</tr>
<tr>
<td>Sum of Total</td>
<td>2662</td>
<td>2623</td>
<td>3169</td>
<td>3113</td>
</tr>
<tr>
<td>Times</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
First Impressions of OS Styles Affect Usability

**Figure 3**

<table>
<thead>
<tr>
<th></th>
<th>GUI Yes</th>
<th>GUI No</th>
<th>CLI Yes</th>
<th>CLI No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enjoyment</td>
<td>18</td>
<td>12</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Accuracy</td>
<td>17</td>
<td>13</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Efficiency</td>
<td>16</td>
<td>14</td>
<td>12</td>
<td>18</td>
</tr>
</tbody>
</table>
First Impressions of OS Styles Affect Usability

Figure 4

![Bar Chart]

Mean Total Time (Seconds)

<table>
<thead>
<tr>
<th>TYPE</th>
<th>Time (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>210</td>
</tr>
<tr>
<td>MacOs</td>
<td>220</td>
</tr>
<tr>
<td>Dove</td>
<td>180</td>
</tr>
<tr>
<td>console</td>
<td>170</td>
</tr>
</tbody>
</table>
First Impressions of OS Styles Affect Usability

Figure 5
First Impressions of OS Styles Affect Usability

Figure 6

Perceived Expertise on the Two Interface Styles

--- GUI --- CLI

Number of Participants

Perceived Expertise

- 30 -